GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 22:57:59 ; Search time 16228 Seconds
(without alignments)
7058.672 Million cell updates/sec

Title: US-10-052-482-162
Perfect score: 2364
Sequence: 1 atggctgtgtggatacaagc.....tcccgcacgcacaatcgtga 2364

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

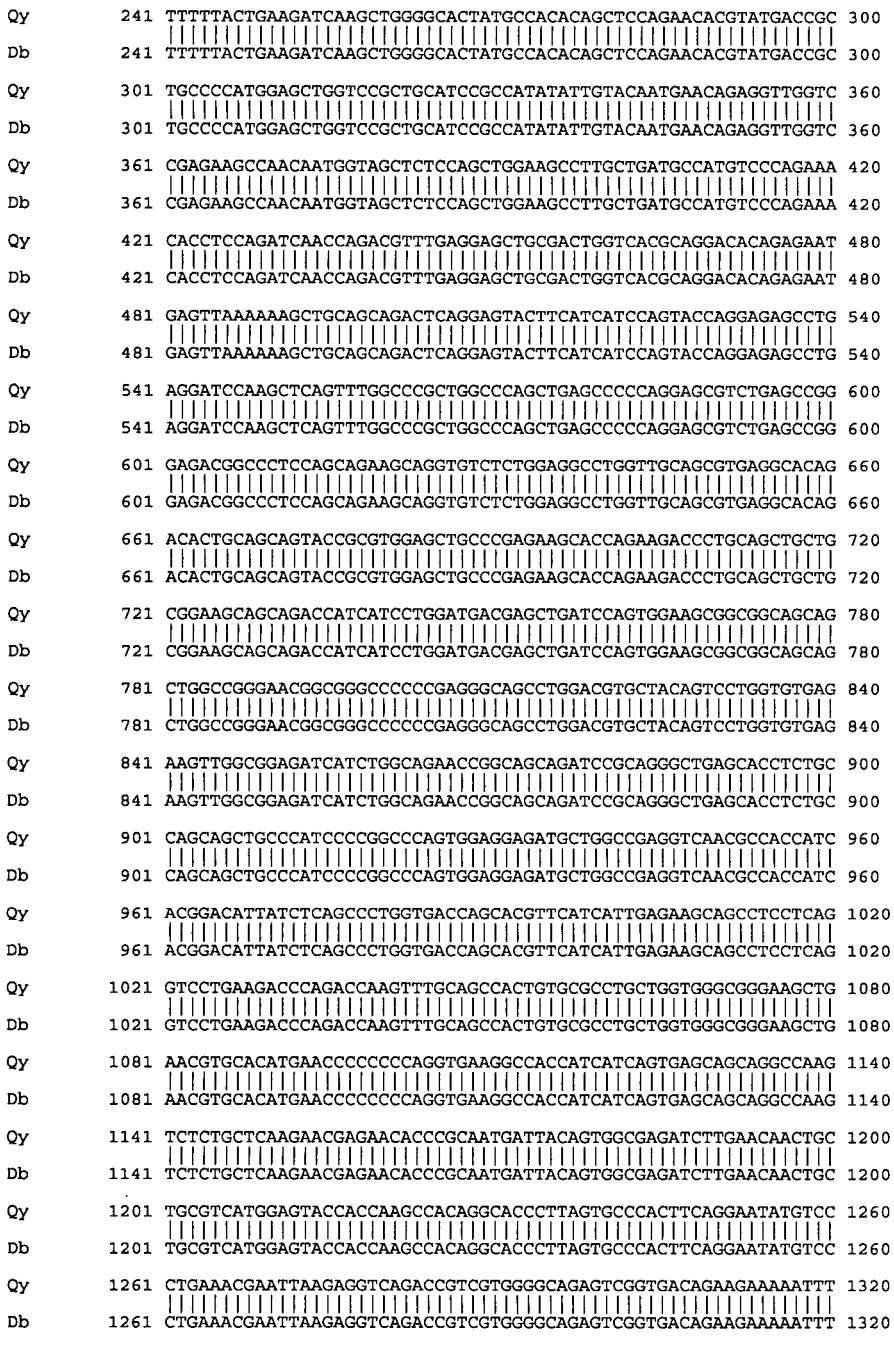
Result No.	Score	% Query Match	Length	DB	ID	Description
1	2364	100.0	2364	6	AX695547	AX695547 Sequence
2	2364	100.0	2782	6	AR360019	AR360019 Sequence
3	2364	100.0	2782	6	AX622837	AX622837 Sequence
4	2364	100.0	2782	6	AX695546	AX695546 Sequence
5	2364	100.0	2782	6	AX777761	AX777761 Sequence
6	2364	100.0	2782	9	HSU47686	U47686 Human signa
7	2356	99.7	2364	6	CQ815791	CQ815791 Sequence
8	2356	99.7	2716	9	HSU48730	U48730 Homo sapien
9	2356	99.7	5167	9	BC065227	BC065227 Homo sapi
10	2096.8	88.7	2414	4	AF135123	AF135123 Sus scrof
11	2079.2	88.0	2485	4	BTA237933	AJ237933 Bos tauru
12	1991.2	84.2	2671	10	RNSTAT5BG	X97541 R.norvegicu
13	1981.6	83.8	4893	10	BC024319	BC024319 Mus muscu
14	1980	83.8	2940	10	AY040231	AY040231 Mus muscu
15	1980	83.8	2940	10	AY044901	AY044901 Mus muscu
16	1980	83.8	2940	10	AY044902	AY044902 Mus muscu
17	1980	83.8	2940	10	AY044903	AY044903 Mus muscu
18	1978.4	83.7	2940	6	AR359993	AR359993 Sequence
19	1978.4	83.7	2940	6	AX622811	AX622811 Sequence
20	1978.4	83.7	2940	6	AX777735	AX777735 Sequence
21	1978.4	83.7	2940	10	AY042906	AY042906 Mus muscu
22	1978.4	83.7	5236	10	MMU21110	U21110 Mus musculu
23	1976.8	83.6	2361	6	CQ815795	CQ815795 Sequence
24	1976.8	83.6	2361	6	AX695544	AX695544 Sequence
25	1976.8	83.6	2574	6	AX695543	AX695543 Sequence
26	1976.8	83.6	2574	10	MMSTAT5B	Z48539 M.musculus
27	1975.2	83.6	2615	10	RNSTAT5B	X91988 R.norvegicu
28	1896.4	80.2	3778	9	BC027036	BC027036 Homo sapi
29	1894.8	80.2	2385	6	CQ815789	CQ815789 Sequence
30	1894.8	80.2	2385	6	I40052	I40052 Sequence 1
31	1894.8	80.2	3120	6	A84439	A84439 Sequence 54
32	1894.8	80.2	3120	6	CQ829963	CQ829963 Sequence
33	1894.8	80.2	3120	6	AR279910	AR279910 Sequence
34	1894.8	80.2	3120	6	AX427338	AX427338 Sequence
35	1894.8	80.2	3120	6	BD082532	BD082532 A method
36	1894.8	80.2	3138	6	A84463	A84463 Sequence 78
37	1894.8	80.2	3138	6	CQ829987	CQ829987 Sequence
38	1894.8	80.2	3138	6	AR279922	AR279922 Sequence
39	1894.8	80.2	3138	6	AX427362	AX427362 Sequence
40	1894.8	80.2	3138	6	BD082544	BD082544 A method
41	1894.8	80.2	3714	6	AR121652	AR121652 Sequence
42	1894.8	80.2	3714	6	AR380925	AR380925 Sequence
43	1894.8	80.2	3714	9	HUMSTAT5A	L41142 Homo sapien
44	1894.8	80.2	4279	6	AR559330	AR559330 Sequence
45	1894.8	80.2	4279	9	HSU43185	U43185 Human signa

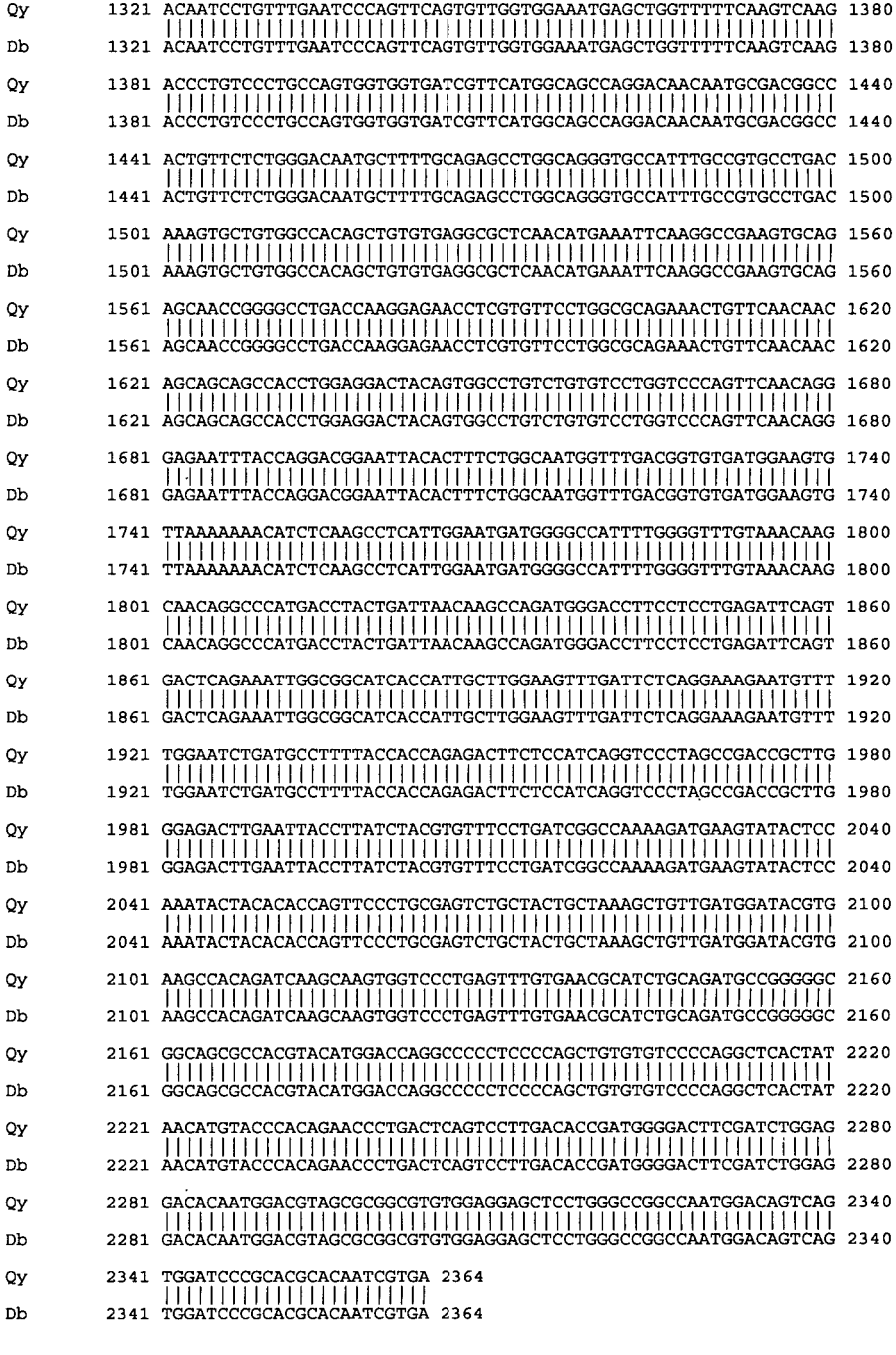
ALIGNMENTS

RESULT 1
AX695547
LOCUS AX695547 2364 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 1174 from Patent WO03008583.
ACCESSION AX695547
VERSION AX695547.1 GI:29418699
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Morris,D.W. and Engelhard,E.K.
TITLE Novel compositions and methods for cancer
JOURNAL Patent: WO 03008583-A 1174 30-JAN-2003;
Sagres Discovery (US)
FEATURES Location/Qualifiers
source 1. .2364
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 2364; DB 6; Length 2364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
```


Qy 241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
Db 241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
Qy 301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
Db 301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
Qy 361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
Db 361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
Qy 421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
Db 421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
Qy 481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
Db 481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
Qy 541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
Db 541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
Qy 601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
Db 601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
Qy 661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
Db 661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
Qy 721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
Db 721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
Qy 781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
Db 781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
Qy 841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
Db 841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
Qy 901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
Db 901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
Qy 961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
Db 961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
Qy 1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
Db 1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
Qy 1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
Db 1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
Qy 1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
Db 1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
Qy 1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
Db 1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
Qy 1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
Db 1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320


Qy 1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
Db 1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
Qy 1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
Db 1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
Qy 1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
Db 1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
Qy 1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
Db 1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
Qy 1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
Db 1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
Qy 1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
Db 1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
Qy 1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
Db 1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
Qy 1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
Db 1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
Qy 1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
Db 1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
Qy 1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
Db 1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
Qy 1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
Db 1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
Qy 1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
Db 1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
Qy 2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
Db 2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
Qy 2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
Db 2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
Qy 2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
Db 2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
Qy 2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
Db 2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
Qy 2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
Db 2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
Qy 2341 TGGATCCCGCACGCACAATCGTGA 2364
Db 2341 TGGATCCCGCACGCACAATCGTGA 2364

```
RESULT 2
AR360019
LOCUS       AR360019               2782 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 158 from patent US 6596036.
ACCESSION   AR360019
VERSION     AR360019.1  GI:33766863
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2782)
  AUTHORS   Derksen,G.C.H., Van Beek,T.A., De Groot,A. and Capelle,A.
  TITLE     Method for the production of a dye preparation based on madder root
  JOURNAL   Patent: US 6596036-A 158 22-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .2782
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             100.0%;  Score 2364;  DB 6;  Length 2782;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2364;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
              |||||||||||||||||||||||||
Db       2487 TGGATCCCGCACGCACAATCGTGA 2510


RESULT 3
AX622837
LOCUS       AX622837               2782 bp    DNA     linear   PAT 20-FEB-2003
DEFINITION  Sequence 158 from Patent WO0224867.
ACCESSION   AX622837
VERSION     AX622837.1  GI:28450816
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pedersen,F.S., Sorensen,A.B., Hernandez,J.M., Nielsen,A.A. and
            Moving,H.O.
  TITLE     Novel compositions and methods for lymphoma and leukemia
  JOURNAL   Patent: WO 0224867-A 158 28-MAR-2002;
            Aarhus University (DK)
FEATURES             Location/Qualifiers
     source          1. .2782
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 2364;  DB 6;  Length 2782;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2364;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy        481 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346
```

```
Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186

Qy       2041 AAATACTACACCACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 AAATACTACACCACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
             ||||||||||||||||||||||||
Db       2487 TGGATCCCGCACGCACAATCGTGA 2510
```

RESULT 4
AX695546
LOCUS AX695546 2782 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 1173 from Patent WO03008583.
ACCESSION AX695546
VERSION AX695546.1 GI:29418698
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Morris,D.W. and Engelhard,E.K.
TITLE Novel compositions and methods for cancer
JOURNAL Patent: WO 03008583-A 1173 30-JAN-2003;
Sagres Discovery (US)
FEATURES Location/Qualifiers
source 1. .2782
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 2364; DB 6; Length 2782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy        481 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
```

```
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660

Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720

Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780

Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840

Db        927 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900

Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960

Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020

Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080

Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140

Db       1227 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200

Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260

Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320

Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380

Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440

Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500

Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560

Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620

Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680

Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740

Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800

Db       1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860

Db       1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920

Db       2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980

Db       2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040

Db       2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100

Db       2187 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160

Db       2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220

Db       2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280

Db       2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340

Db       2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364

Db       2487 TGGATCCCGCACGCACAATCGTGA 2510
```

```
RESULT 5
AX777761
LOCUS       AX777761               2782 bp    DNA     linear   PAT 14-JUL-2003
DEFINITION  Sequence 158 from Patent WO03027276.
ACCESSION   AX777761
VERSION     AX777761.1  GI:32694773
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pedersen,F.S., Sorensen,A.B., Hernandez,J.M., Nielsen,A.A. and
            Moving,H.
  TITLE     Novel compositions and methods for diagnosis and treatment of
            lymphoma and leukemia
  JOURNAL   Patent: WO 03027276-A 158 03-APR-2003;
            UNIVERSITY OF AARHUS (DK)
FEATURES             Location/Qualifiers
     source          1. .2782
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
```

/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 2364; DB 6; Length 2782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy        481 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246
```

```
Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
             ||||||||||||||||||||||||
Db       2487 TGGATCCCGCACGCACAATCGTGA 2510
```

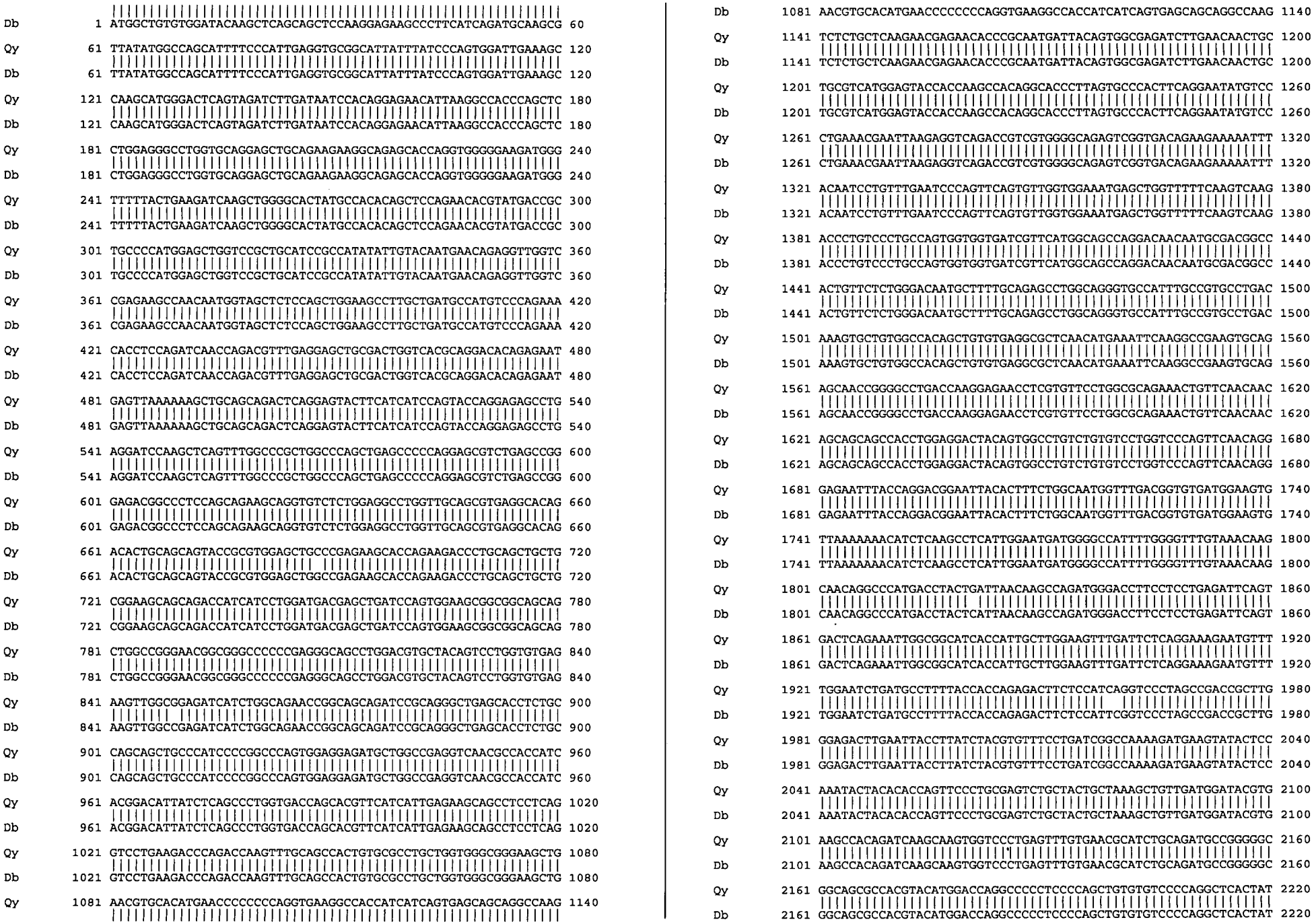
RESULT 6
HSU47686
LOCUS HSU47686 2782 bp mRNA linear . PRI 23-MAY-1996
DEFINITION Human signal transducer and activator of transcription Stat5B mRNA, complete cds.
ACCESSION U47686
VERSION U47686.1 GI:1330323
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2782)
AUTHORS Lin,J.X., Mietz,J., Modi,W.S., John,S. and Leonard,W.J.
TITLE Cloning of human Stat5B. Reconstitution of interleukin-2-induced Stat5A and Stat5B DNA binding activity in COS-7 cells
JOURNAL J. Biol. Chem. 271 (18), 10738-10744 (1996)
MEDLINE 96210005
PUBMED 8631883
REFERENCE 2 (bases 1 to 2782)
AUTHORS Lin,J.-X., Mietz,J., Modi,W.S., John,S. and Leonard,W.J.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-1996) Jian-Xin Lin, Lab of Molecular Immunology, NHLBI, NIH, 9000 Rockville Pike, Bldg. 10, Rm. 7N244, Bethesda, MD 20892-1674, USA
FEATURES Location/Qualifiers
source 1. .2782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q11.2"
CDS 147. .2510
/note="STAT protein; is activated by IL-2, IL-7, IL-15, growth hormone, IL-3, GM-CSF, thrombopoietin, prolactin, and erythropoietin; tyrosine 699 phosphorylation is required for activation and dimerization of Stat5B"
/codon_start=1
/product="signal transducer and activator of transcription Stat5B"
/protein_id="AAC50491.1"
/db_xref="GI:1330324"
/translation="MAVWIQAQQLQGEALHQMQALYGQHFPIEVRHYLSQWIESQAWD
SVDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQNTYDRCP
MELVRCIRHILYNEQRLVREANNGSSPAGSLADAMSQKHLQINQTFEELRLVTQDTEN
ELKKLQQTQEYFIIQYQESLRIQAQFGPLAQLSPQERLSRETALQQKQVSLEAWLQRE
AQTLQQYRVELPEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISALVTSTFII
EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQQAKSLLKNENTRNDY
SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKFTILFESQFSV
GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
EALNMKFKAEVQSNRGLTKENLVFLAQKLFNNSSSHLEDYSGLSVSWSQFNRENLPGR
NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI
GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLNYLIYVFPDRPKDEVYSKY
YTPVPCESATAKAVDGYVKPQIKQVVPEFVNASADAGGGSATYMDQAPSPAVCPQAHY
NMYPQNPDSVLDTDGDFDLEDTMDVARRVEELLGRPMDSQWIPHAQS"
ORIGIN

```
  Query Match           100.0%;  Score 2364;  DB 9;  Length 2782;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2364;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186
```

```
Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
              ||||||||||||||||||||||||
Db       2487 TGGATCCCGCACGCACAATCGTGA 2510
```

```
RESULT 7
CQ815791
LOCUS       CQ815791                2364 bp    DNA     linear   PAT 03-JUN-2004
DEFINITION  Sequence 3 from Patent WO2004042040.
ACCESSION   CQ815791
VERSION     CQ815791.1  GI:48144323
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Nakauchi,H. and Iwama,A.
  TITLE     Expansion agents for stem cells
  JOURNAL   Patent: WO 2004042040-A 3 21-MAY-2004;
            ReproCELL Inc. (JP)
FEATURES             Location/Qualifiers
     source          1. .2364
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .2364
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAG32954.1"
                     /db_xref="GI:48144324"
                     /translation="MAVWIQAQQLQGEALHQMQALYGQHFPIEVRHYLSQWIESQAWD
                     SVDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQNTYDRCP
                     MELVRCIRHILYNEQRLVREANNGSSPAGSLADAMSQKHLQINQTFEELRLVTQDTEN
                     ELKKLQQTQEYFIIQYQESLRIQAQFGPLAQLSPQERLSRETALQQKQVSLEAWLQRE
                     AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
                     SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISALVTSTFII
                     EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQQAKSLLLKNENTRNDY
                     SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKFTILFESQFSV
                     GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
                     EALNMKFKAEVQSNRGLTKENLVFLAQKLFNNSSSHLEDYSGLSVSWSQFNRENLPGR
                     NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI
                     GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLNYLIYVFPDRPKDEVYSKY
                     YTPVPCESATAKAVDGYVKPQIKQVVPEFVNASADAGGGSATYMDQAPSPAVCPQAHY
                     NMYPQNPDSVLDTDGDFDLEDTMDVARRVEELLGRPMDSQWIPHAQS"
ORIGIN

  Query Match           99.7%;  Score 2356;  DB 6;  Length 2364;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 2359;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
```



```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        661 ACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGTTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1801 CAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              ||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCGACCGCTTG 1980

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
```

```
Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
             |||||||||||||||||||||||||
Db       2341 TGGATCCCGCACGCACAATCGTGA 2364


RESULT 8
HSU48730
LOCUS       HSU48730                2716 bp    mRNA    linear   PRI 16-FEB-2000
DEFINITION  Homo sapiens transcription factor Stat5b (stat5b) mRNA, complete
            cds.
ACCESSION   U48730
VERSION     U48730.2  GI:6981725
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2716)
  AUTHORS   Silva,C.M., Lu,H. and Day,R.N.
  TITLE     Characterization and cloning of STAT5 from IM-9 cells and its
            activation by growth hormone
  JOURNAL   Mol. Endocrinol. 10 (5), 508-518 (1996)
  MEDLINE   96311205
   PUBMED   8732682
REFERENCE   2  (bases 1 to 2716)
  AUTHORS   Silva,C.M. and Lu,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-FEB-1996) Internal Medicine, University of Virginia,
            Box 511, Charlottesville, VA 22908, USA
REFERENCE   3  (bases 1 to 2716)
  AUTHORS   Silva,C.M. and Lu,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2000) Internal Medicine, University of Virginia,
            Box 511, Charlottesville, VA 22908, USA
  REMARK    Sequence update by submitter
COMMENT     On Feb 16, 2000 this sequence version replaced gi:1216480.
FEATURES             Location/Qualifiers
     source          1. .2716
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_line="IM-9"
                     /cell_type="lymphocyte"
     gene            1. .2716
                     /gene="stat5b"
     CDS             12. .2375
                     /gene="stat5b"
                     /function="signal transducer and activator of
                     transcription"
                     /codon_start=1
                     /product="transcription factor Stat5b"
                     /protein_id="AAC50485.2"
                     /db_xref="GI:6981726"
                     /translation="MAVWIQAQQLQGEALHQMQALYGQHFPIEVRHYLSQWIESQAWD
                     SVDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQNTYDRCP
                     MELVRCIRHILYNEQRLVREANNGSSPAGSLADAMSQKHLQINQTFEELRLVTQDTEN
                     ELKKLQQTQEYFIIQYQESLRIQAQFGPLAQLSPQERLSRETALQQKQVSLEAWLQRE
                     AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
                     SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISALVTSTFII
                     EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQQAKSLLKNENTRNDY
                     SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKFTILFESQFSV
                     GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
                     EALNMKFKAEVQSNRGLTKENLVFLAQKLFNNSSSHLEDYSGLSVSWSQFNRENLPGR
                     NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI
                     GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLNYLIYVFPDRPKDEVYSKY
                     YTPVPCESATAKAVDGYVKPQIKQVVPEFVNASADAGGGSATYMDQAPSPAVCPQAHY
                     NMYPQNPDSVLDTDGDFDLEDTMDVARRVEELLGRPMDSQWIPHAQS"
ORIGIN

  Query Match           99.7%;  Score 2356;  DB 9;  Length 2716;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 2359;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db         12 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAGGCG 71

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 131

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 191

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 251

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 311

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 371

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 431

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 491

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 551

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 611

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 671

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        672 ACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 731

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 791

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 851

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 AAGTTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 911

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 971
```

```
Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1031

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1032 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1091

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1151

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1152 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1211

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1212 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1271

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1272 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1331

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1332 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1391

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1392 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1451

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1452 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1511

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1512 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1571

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1572 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1631

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1632 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1691

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1692 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1751

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1752 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1811

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1812 CAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1871

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1872 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1931

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1932 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCCGGTCCCTAGCCGACCGCTTG 1991

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1992 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2051

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2052 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2111

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2112 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2171

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2172 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2231

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2232 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2291

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2292 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2351

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
              ||||||||||||||||||||||||
Db       2352 TGGATCCCGCACGCACAATCGTGA 2375
```

RESULT 9
BC065227

LOCUS	BC065227 5167 bp mRNA linear PRI 12-OCT-2004
DEFINITION	Homo sapiens cDNA clone MGC:74606 IMAGE:5553426, complete cds.
ACCESSION	BC065227
VERSION	BC065227.1 GI:40807019
KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 5167)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 5167)
AUTHORS	Director MGC Project.
TITLE	Direct Submission
JOURNAL	Submitted (12-JAN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: Lou Staudt cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),

Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 139 Row: b Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 42519913.

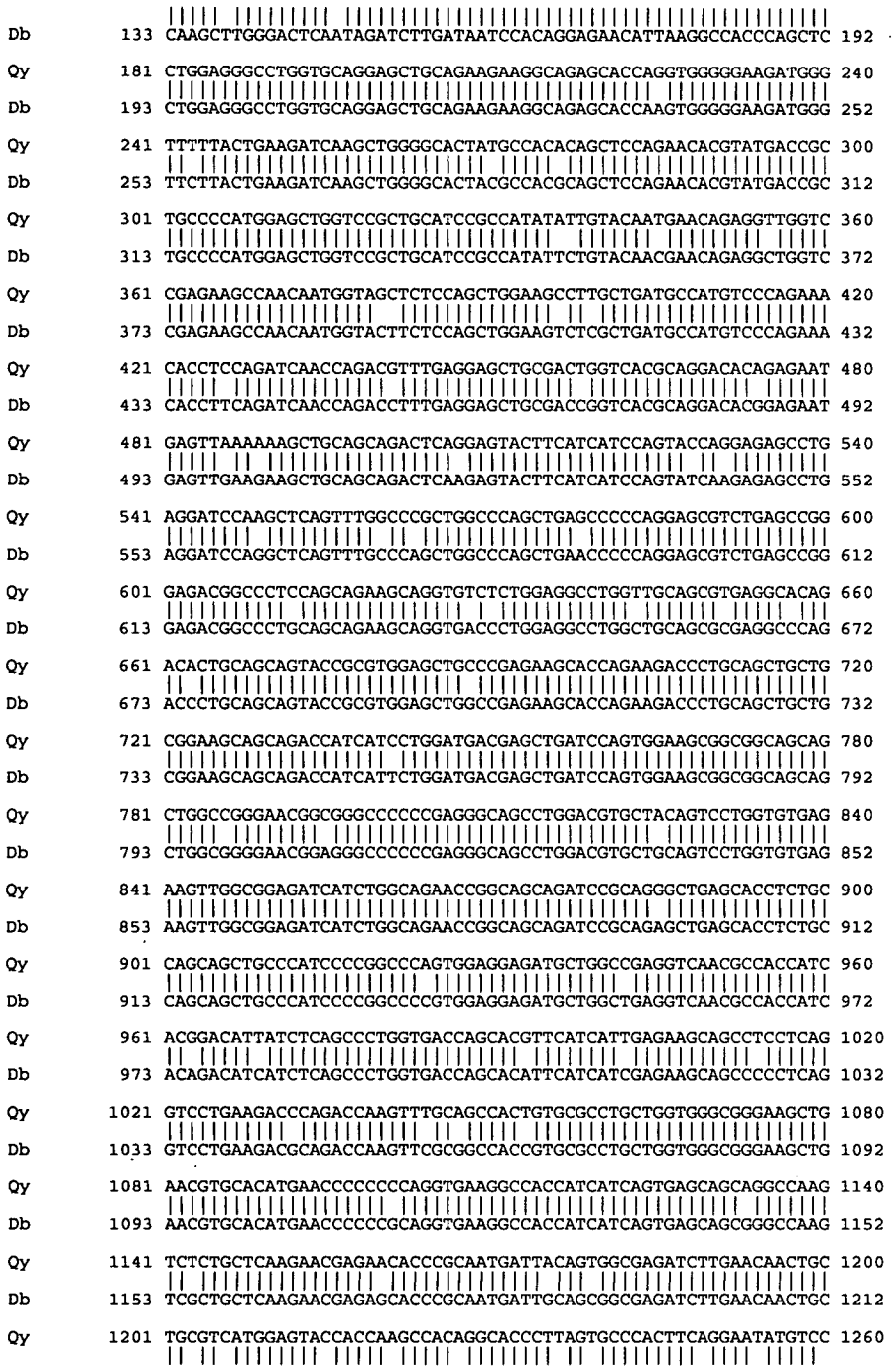
FEATURES Location/Qualifiers
source 1. .5167
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:74606 IMAGE:5553426"
/tissue_type="Lymph, lymphoma"
/clone_lib="NIH_MGC_85"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
CDS 170. .2533
/codon_start=1
/product="Unknown (protein for MGC:74606)"
/protein_id="AAH65227.1"
/db_xref="GI:40807020"
/translation="MAVWIQAQQLQGEALHQMQALYGQHFPIEVRHYLSQWIESQAWD
SVDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQNTYDRCP
MELVRCIRHILYNEQRLVREANNGSSPAGSLADAMSQKHLQINQTFEELRLVTQDTEN
ELKKLQQTQEYFIIQYQESLRIQAQFGPLAQLSPQERLSRETALQQKQVSLEAWLQRE
AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISALVTSTFII
EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQQAKSLLKNENTRNDY
SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKFTILFESQFSV
GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
EALNMKFKAEVQSNRGLTKENLVFLAQKLFNNSSSHLEDYSGLSVSWSQFNRENLPGR
NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI
GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLNYLIYVFPDRPKDEVYSKY
YTPVPCESATAKAVDGYVKPQIKQVVPEFVNASADAGGGSATYMDQAPSPAVCPQAHY
NMYPQNPDSVLDTDGDFDLEDTMDVARRVEELLGRPMDSQWIPHAQS"

ORIGIN

Query Match 99.7%; Score 2356; DB 9; Length 5167;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2359; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 229

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 289

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 349

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 409

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 469

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 529

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 589

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        590 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 649

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        650 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 709

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        710 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 769

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        770 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 829

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        830 ACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 889

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        890 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 949

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        950 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 1009

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 AAGTTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1069

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1070 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1129

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1130 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1189

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1190 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1249

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1250 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1309

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1310 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1369

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1370 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1429

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1430 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1489

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1490 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1549

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1550 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1609

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1610 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1669

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1670 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1729

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1730 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1789

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1790 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1849

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1850 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1909

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1910 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1969

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1970 CAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2029

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2030 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2089

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db       2090 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCGACCGCTTG 2149

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2150 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2209

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2210 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2269

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2270 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2329

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2330 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2389

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2390 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2449

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2450 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2509

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
              |||||||||||||||||||||||||
Db       2510 TGGATCCCGCACGCACAATCGTGA 2533


RESULT 10
AF135123
LOCUS       AF135123               2414 bp    mRNA    linear   MAM 01-AUG-1999
DEFINITION  Sus scrofa signal transducer and activator of transcription 5b
            (Stat5b) mRNA, complete cds.
ACCESSION   AF135123
VERSION     AF135123.1  GI:5669104
KEYWORDS    .
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 2414)
  AUTHORS   Palin,M.F., Beaudry,D., Roberge,C. and Farmer,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-1999) Agriculture and Agri-Food Canada, Dairy and
            Swine Research and Development Centre, P.O. Box 90, 2000 Route 108
            East, Lennoxville, Quebec J1M 1Z3, Canada
FEATURES             Location/Qualifiers
     source          1. .2414
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /strain="breed Large White"
                     /db_xref="taxon:9823"
                     /tissue_type="mammary gland"
     gene            1. .2414
                     /gene="Stat5b"
     CDS             13. .2376
                     /gene="Stat5b"
                     /note="mammary gland factor; transcription factor"
                     /codon_start=1
                     /product="signal transducer and activator of transcription
                     5b"
                     /protein_id="AAD46164.1"
                     /db_xref="GI:5669105"
                     /translation="MAVWIQAQQLQGDALHQMQALYGQHFPIEVRHYLSQWIESQAWD
                     SIDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQNTYDRCP
                     MELVRCIRHILYNEQRLVREANNGTSPAGSLADAMSQKHLQINQTFEELRPVTQDTEN
                     ELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLNPQERLSRETALQQKQVTLEAWLQRE
                     AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
                     SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISALVTSTFII
                     EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQRAKSLLKNESTRNDC
                     SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKFTILFESQFSV
                     GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
                     EALNMKFKAEVQSNRGLTKENLVFLAQKLFNSSSSHLEDYSGMSVSWSQFNRENLPGR
                     NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI
                     GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLNYLIYVFPDRPKDEVYSKY
                     YTPVPCEPATAKAVDGYVKPQIKQVVPEFVSASSDSAGGNATYMDQAPSPAVCPQAHY
                     SIYPQNPDPVLDNDGDFDLDDTMDVARRVEELLGRPMDSQWIPHAQS"
     variation       69
                     /gene="Stat5b"
                     /replace="g"
     variation       470
                     /gene="Stat5b"
                     /replace="t"
     variation       1145
                     /gene="Stat5b"
                     /replace="a"
     variation       1187
                     /gene="Stat5b"
                     /replace="a"
     variation       1273
                     /gene="Stat5b"
                     /replace="t"
ORIGIN

  Query Match             88.7%;  Score 2096.8;  DB 4;  Length 2414;
  Best Local Similarity   92.9%;  Pred. No. 0;
  Matches 2197;  Conservative    0;  Mismatches  167;  Indels    0;  Gaps    0;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              |||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db         13 ATGGCTGTGTGGATACAAGCCCAGCAGCTCCAAGGAGATGCCCTTCATCAGATGCAAGCA 72

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
                | ||||| |||||||| ||||| ||||||||||||||||||||||||||||||||||
Db         73 CTGTATGGTCAGCATTTCCCCATCGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 132

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
```


Db 133 CAAGCTTGGGACTCAATAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 192
Qy 181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
Db 193 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAAGTGGGGGAAGATGGG 252
Qy 241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
Db 253 TTCTTACTGAAGATCAAGCTGGGGCACTACGCCACGCAGCTCCAGAACACGTATGACCGC 312
Qy 301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
Db 313 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATTCTGTACAACGAACAGAGGCTGGTC 372
Qy 361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
Db 373 CGAGAAGCCAACAATGGTACTTCTCCAGCTGGAAGTCTCGCTGATGCCATGTCCCAGAAA 432
Qy 421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
Db 433 CACCTTCAGATCAACCAGACCTTTGAGGAGCTGCGACCGGTCACGCAGGACACGGAGAAT 492
Qy 481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
Db 493 GAGTTGAAGAAGCTGCAGCAGACTCAAGAGTACTTCATCATCCAGTATCAAGAGAGCCTG 552
Qy 541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
Db 553 AGGATCCAGGCTCAGTTTGCCCAGCTGGCCCAGCTGAACCCCCAGGAGCGTCTGAGCCGG 612
Qy 601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
Db 613 GAGACGGCCCTGCAGCAGAAGCAGGTGACCCTGGAGGCCTGGCTGCAGCGCGAGGCCCAG 672
Qy 661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
Db 673 ACCCTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 732
Qy 721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
Db 733 CGGAAGCAGCAGACCATCATTCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 792
Qy 781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
Db 793 CTGGCGGGGAACGGAGGGCCCCCCGAGGGCAGCCTGGACGTGCTGCAGTCCTGGTGTGAG 852
Qy 841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
Db 853 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGAGCTGAGCACCTCTGC 912
Qy 901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
Db 913 CAGCAGCTGCCCATCCCCGGCCCCGTGGAGGAGATGCTGGCTGAGGTCAACGCCACCATC 972
Qy 961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
Db 973 ACAGACATCATCTCAGCCCTGGTGACCAGCACATTCATCATCGAGAAGCAGCCCCCTCAG 1032
Qy 1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
Db 1033 GTCCTGAAGACGCAGACCAAGTTCGCGGCCACCGTGCGCCTGCTGGTGGGCGGGAAGCTG 1092
Qy 1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
Db 1093 AACGTGCACATGAACCCCCCGCAGGTGAAGGCCACCATCATCAGTGAGCAGCGGGCCAAG 1152
Qy 1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
Db 1153 TCGCTGCTCAAGAACGAGAGCACCCGCAATGATTGCAGCGGCGAGATCTTGAACAACTGC 1212
Qy 1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
Db 1213 TGTGTGATGGAGTATCACCAGGCCACCGGCACCCTCAGCGCCCACTTCCGGAACATGTCG 1272
Qy 1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
Db 1273 CTGAAACGAATTAAGAGGTCTGACCGTCGTGGAGCAGAGTCAGTGACAGAAGAAAAATTT 1332
Qy 1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
Db 1333 ACAATCTTATTTGAATCACAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1392
Qy 1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
Db 1393 ACCCTGTCCCTCCCAGTGGTGGTGATTGTTCACGGCAGCCAGGACAACAACGCGACGGCC 1452
Qy 1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
Db 1453 ACCGTCCTCTGGGACAATGCTTTCGCAGAGCCCGGCAGGGTGCCCTTTGCGGTGCCTGAC 1512
Qy 1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
Db 1513 AAAGTCCTGTGGCCGCAGCTGTGCGAGGCGCTCAACATGAAATTCAAGGCCGAGGTGCAG 1572
Qy 1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
Db 1573 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAGCTCTTCAACAGC 1632
Qy 1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
Db 1633 AGCAGCAGCCATCTGGAGGACTACAGCGGCATGTCTGTGTCCTGGTCCCAGTTCAACAGG 1692
Qy 1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
Db 1693 GAAAATTTACCAGGACGGAATTACACTTTCTGGCAGTGGTTTGATGGCGTGATGGAAGTG 1752
Qy 1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
Db 1753 TTAAAAAAACATCTCAAGCCTCATTGGAACGATGGGGCTATTTTGGGTTTCGTGAACAAG 1812
Qy 1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
Db 1813 CAACAGGCCCATGACCTACTCATTAACAAGCCCGACGGGACTTTCCTGCTGAGGTTCAGT 1872
Qy 1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
Db 1873 GACTCTGAAATTGGTGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1932
Qy 1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
Db 1933 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCTATCCGGTCCCTGGCAGACCGCTTG 1992
Qy 1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
Db 1993 GGAGACCTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAGGATGAAGTGTACTCC 2052
Qy 2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
Db 2053 AAATATTACACCCCAGTTCCATGCGAGCCTGCCACCGCTAAAGCAGTGGACGGATACGTG 2112
Qy 2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
Db 2113 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAGCGCCTCTTCAGACTCTGCCGGG 2172
Qy 2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
Db 2173 GGCAATGCCACCTACATGGACCAGGCCCCCTCCCCAGCCGTGTGCCCCCAGGCTCATTAC 2232
Qy 2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
Db 2233 AGCATATACCCACAGAACCCCGACCCCGTCCTTGACAACGATGGGGACTTTGATCTGGAC 2292
Qy 2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
Db 2293 GACACGATGGACGTGGCGCGTCGTGTGGAGGAGCTCTTGGGCCGACCTATGGACAGTCAG 2352

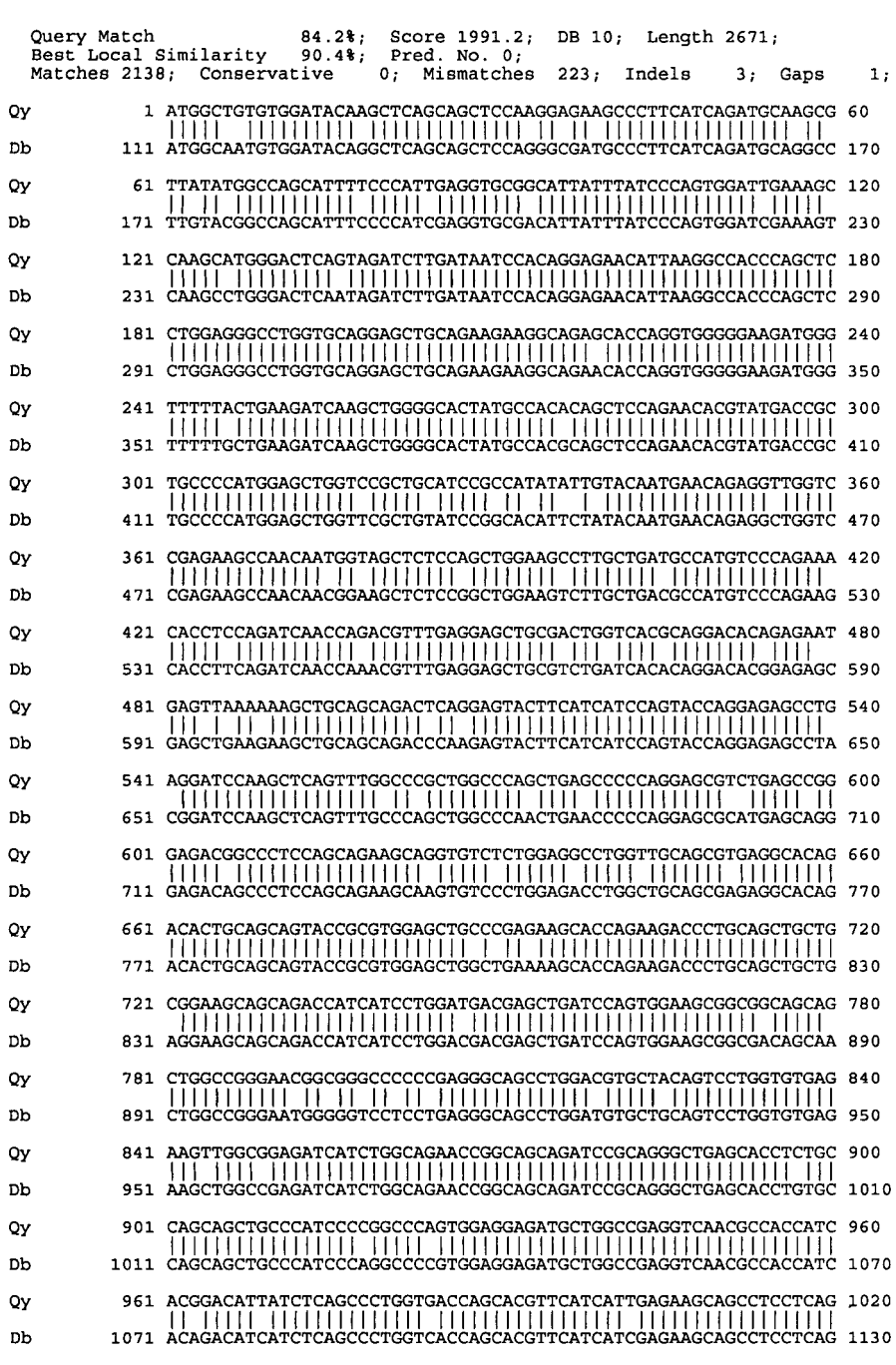
```
Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
             |||||||| ||||| || ||||||
Db       2353 TGGATCCCTCACGCGCAGTCGTGA 2376


RESULT 11
BTA237933
LOCUS       BTA237933               2485 bp    mRNA    linear   MAM 15-JUN-2000
DEFINITION  Bos taurus mRNA for mammary gland factor STAT5B.
ACCESSION   AJ237933
VERSION     AJ237933.1  GI:5689666
KEYWORDS    Mammary Gland Factor STAT5B; stat5b gene.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE   1
  AUTHORS   Seyfert,H.M., Pitra,C., Meyer,L., Brunner,R.M., Wheeler,T.T.,
            Molenaar,A., McCracken,J.Y., Herrmann,J., Thiesen,H.J. and
            Schwerin,M.
  TITLE     Molecular characterization of STAT5A- and STAT5B-encoding genes
            reveals extended intragenic sequence homogeneity in cattle and
            mouse and different degrees of divergent evolution of various
            domains
  JOURNAL   J. Mol. Evol. 50 (6), 550-561 (2000)
  MEDLINE   20296816
   PUBMED   10835485
REFERENCE   2  (bases 1 to 2485)
  AUTHORS   Seyfert,H.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-1999) Seyfert H.M., Molecular Biology, Research
            Institute for the Biology of Farm Animals, Wilhelm-Stahl-Allee-2,
            18196 Dummerstorf, GERMANY
FEATURES             Location/Qualifiers
     source          1. .2485
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="mammary gland"
                     /dev_stage="lactating"
     gene            1. .2485
                     /gene="stat5B"
     CDS             2. .2365
                     /gene="stat5B"
                     /codon_start=1
                     /product="mammary gland factor STAT5B"
                     /protein_id="CAB51845.1"
                     /db_xref="GI:5689667"
                     /db_xref="GOA:Q9TUM3"
                     /db_xref="UniProt/Swiss-Prot:Q9TUM3"
                     /translation="MAVWIQAQQLQGDALHQMQALYGQHFPIEVRHYLSQWIEGQAWD
                     SIDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQNTYDRCP
                     MELVRCIRHILYNEQRLVREANNGTSPAGSFADAMSQKHLQINQTFEELRLVTQDTEN
                     ELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLNPQERLSRETALQQKQVSLEAWLQRE
                     AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
                     SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISALVTSTFII
                     EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQQAKSLLKNENTRNDY
                     SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKFTILFESQFSV
                     GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
                     EALNMKFKAEVQSNRGLTKENLVFLAQKLFNSSSSHLEDYNGMSVSWSQFNRENLPGR
                     NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKSDGTFLLRFSDSEI
                     GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLSYLIYVFPDRPKDEVYSKY
                     YTPVPCEPATAKAVDGYVKPQIKQVVPEFVSASADSAGGSATYMDQAPSPAVCPQPHY
                     NMYPQNPDPVLDNDGDFDLDDTIDVARRVEELLGRPMDSQWIPHAQS"
ORIGIN

  Query Match             88.0%;  Score 2079.2;  DB 4;  Length 2485;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches 2186;  Conservative    0;  Mismatches  178;  Indels    0;  Gaps    0;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              |||||||||||||||||| |||||||||||| || ||||||||||||||||||||||
Db          2 ATGGCTGTGTGGATACAAGCCCAGCAGCTCCAAGGCGATGCCCTTCATCAGATGCAAGCG 61

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
               | || |||||||||| |||||||||||| ||||| |||||||||||||||||| ||
Db         62 CTGTACGGCCAGCATTTCCCCATTGAGGTGCGCCATTACTTATCCCAGTGGATTGAAGGC 121

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        122 CAAGCTTGGGACTCAATAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 181

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              |||||||||||||||||||| |||||||||||||||||||||| |||||||||||| |||
Db        182 CTGGAGGGCCTGGTGCAGGAGTTGCAGAAGAAGGCAGAGCACCAAGTGGGGGAAGACGGG 241

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||  |||||||||||||||||||||| |||||||||||||||||||| |||||||||
Db        242 TTCCTACTGAAGATCAAGCTGGGGCACTACGCCACACAGCTCCAGAACACATATGACCGC 301

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              |||||||||||||||| |||||||||||||| |||||||||||||||| |||||
Db        302 TGCCCCATGGAGCTGGTGCGCTGCATCCGCCATATTCTGTACAATGAACAGAGGCTGGTC 361

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              |||||||||||||||||  |||||| |||||||  | || ||||||||||||||||||
Db        362 CGAGAAGCCAACAATGGTACTTCTCCGGCTGGAAGTTTCGCCGATGCCATGTCCCAGAAA 421

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              |||||  || |||||||||| |||||||||||||||||||||| || ||||||||||||
Db        422 CACCTTCAAATCAACCAGACATTTGAGGAGCTGCGACTGGTCACACAAGACACAGAGAAT 481

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||| | || |||||||||||||||||| |||||||| |||||||||| ||||||||||||
Db        482 GAGCTGAAGAAGCTGCAGCAGACTCAAGAGTACTTTATCATCCAGTATCAGGAGAGCCTG 541

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              |||||||| |||||||||| || |||||||||||||| |||||||||| |||||||||
Db        542 AGGATCCAGGCTCAGTTTGCCCAGCTGGCCCAGCTGAACCCCCAGGAGCGACTGAGCCGG 601

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              |||||||||||||||||||||||||||||| |||||||||| ||||||| ||||| |||
Db        602 GAGACGGCCCTCCAGCAGAAGCAGGTGTCCCTGGAGGCCTGGCTGCAGCGCGAGGCCCAG 661

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||·||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        662 ACGCTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 721

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 781

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||| |||||||| |||||||||||||||||||| ||||||||||||||||||||||
Db        782 CTGGCGGGGAACGGAGGGCCCCCCGAGGGCAGCCTGGATGTGCTACAGTCCTGGTGTGAG 841

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||
Db        842 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGAGCCGAGCACCTCTGC 901

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db        902 CAGCAGCTGCCCATCCCCGGCCCCGTGGAGGAGATGCTGGCTGAGGTCAACGCCACCATC 961

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              || ||||| |||||||||||||||||||||| |||||||| |||||||||| ||||||
Db        962 ACTGACATCATCTCAGCCCTGGTGACCAGCACCTTCATCATCGAGAAGCAGCCCCCTCAG 1021

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              |||||||||||||||||||| || ||||| ||||||||||||||||||||||||||||||
Db       1022 GTCCTGAAGACCCAGACCAAGTTCGCGGCCACCGTGCGCCTGCTGGTGGGCGGGAAGCTG 1081

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              || |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
```

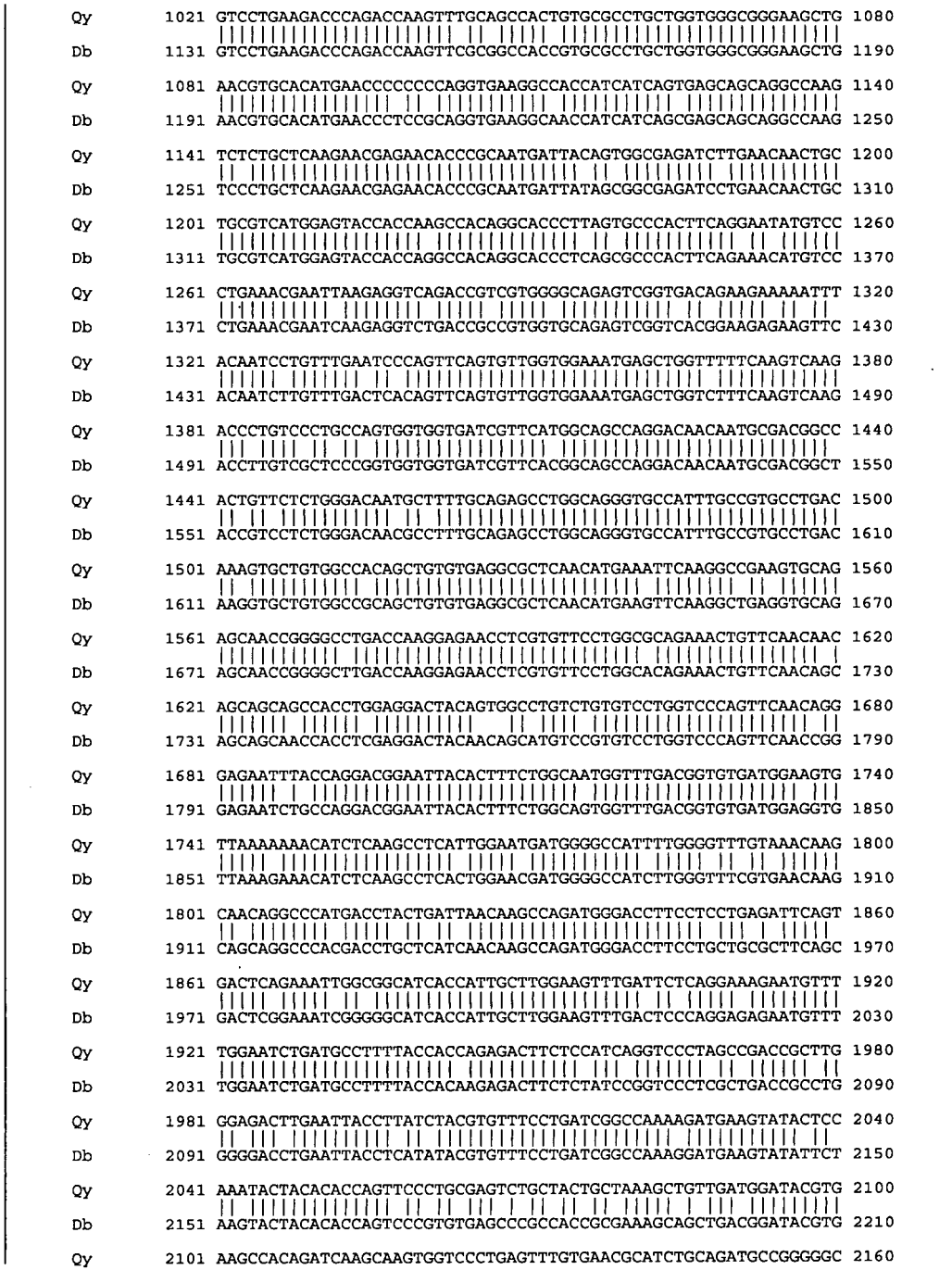
```
Db       1082 AATGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGCGAGCAGCAGGCCAAG 1141

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              || |||||||||||||||||||||||||||||||||||||||||||| |||||| ||||||||||
Db       1142 TCACTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGGGAGATCCTGAACAACTGC 1201

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              || || |||||||||||||| |||||||||||||| || ||||||||| |||| ||||||
Db       1202 TGTGTAATGGAGTACCACCAGGCCACAGGCACCCTGAGCGCCCACTTCCGGAACATGTCC 1261

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||| || ||||| ||||| ||||||||||||||||||||||||||||
Db       1262 CTGAAACGAATTAAGAGATCTGACCGGCGTGGAGCAGAGTCGGTGACAGAAGAAAAATTT 1321

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              || |||||||||||||| |||||||||||||| |||||||||||||| ||||||||||||
Db       1322 ACCATCCTGTTTGAATCACAGTTCAGTGTTGGCGGAAATGAGCTGGTGTTTCAAGTCAAG 1381

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              |||||||| || |||||||||||||||| ||||| |||||||||| ||||||||||||||||
Db       1382 ACCCTGTCTCTCCCAGTGGTGGTGATTGTTCACGGCAGCCAGGATAACAATGCGACGGCC 1441

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              || |||||||||||||| |||||||||||||||||||||||||||| ||||| ||||| |||
Db       1442 ACCGTTCTCTGGGACAACGCTTTTGCAGAGCCTGGCAGGGTGCCGTTTGCAGTGCCCGAC 1501

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||| |||||||| |||||||| |||||||||||||||||||||||||||||| ||||||
Db       1502 AAAGTCCTGTGGCCGCAGCTGTGCGAGGCGCTCAACATGAAATTCAAGGCCGAGGTGCAG 1561

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              |||||||||||||||||||||||||||||||| |||||||||||||||| |||||||||| |
Db       1562 AGCAACCGGGGCCTGACCAAGGAGAACCTGGTGTTCCTGGCGCAGAAGCTGTTCAACAGC 1621

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              |||||||||||||| |||||||||| ||| |||| |||||||||||||||||||||| ||
Db       1622 AGCAGCAGCCACCTCGAGGACTACAACGGCATGTCCGTGTCCTGGTCCCAGTTCAACCGG 1681

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              |||||||||||| |||||||||||||||||||||||| |||||||| ||||| |||||||||
Db       1682 GAGAATTTACCCGGACGGAATTACACTTTCTGGCAGTGGTTTGATGGTGTTATGGAAGTG 1741

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||| |||||||||||||||||||||||||||||| ||||| || ||||| || || ||||||
Db       1742 TTAAAGAAACATCTCAAGCCTCATTGGAATGACGGGGCTATCTTGGGTTTCGTGAACAAG 1801

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              |||||||||| ||||| || || |||||| |||| |||||||||| ||||| ||||||
Db       1802 CAACAGGCCCACGACCTGCTTATCAACAAGTCAGACGGGACCTTCCTGCTGAGGTTCAGT 1861

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||| ||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 GACTCTGAAATCGGTGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1921

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              |||||||||||||||||||||||||||||||| ||||||||| ||||||| ||||||||||||
Db       1922 TGGAATCTGATGCCTTTTACCACCAGAGATTTCTCCATCCGGTCCCTGGCCGACCGCTTG 1981

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              |||||| ||| |||||||||||||||||||||||| ||||| |||||||||||| ||||||
Db       1982 GGAGACCTGAGTTACCTTATCTACGTGTTTCCTGACCGGCCCAAAGATGAAGTGTACTCC 2041

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              || || ||||| |||||||| |||||| |||| |||||||||| || || || ||||||
Db       2042 AAGTATTACACCCCAGTTCCATGCGAGCCTGCCACTGCTAAAGCAGTGGACGGGTACGTG 2101

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              |||||||||||||||| |||||||||||||||||||| ||| |||||||| | |  |||
Db       2102 AAGCCACAGATCAAGCAGGTGGTCCCTGAGTTTGTGAGCGCCTCTGCAGACTCTGCCGGC 2161

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||| ||||| |||||||||||||| |||||||||| ||||| |||||| |||||||||
Db       2162 GGCAGTGCCACCTACATGGACCAGGCTCCCTCCCCAGCCGTGTGCCCCCAGCCTCACTAT 2221

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
              |||||||||||||||||||| ||| | |||||||||| |||||| ||||||||||||||||
Db       2222 AACATGTACCCACAGAACCCCGACCCCGTCCTTGACAACGATGGTGACTTCGATCTGGAC 2281

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
              ||||| || ||||| |||||||||||||||||||| ||||| || ||||||||||||
Db       2282 GACACGATTGACGTGGCGCGGCGTGTGGAGGAGCTCCTAGGCCGACCCATGGACAGTCAG 2341

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
              |||||||| |||||||| ||||||
Db       2342 TGGATCCCCCACGCACAGTCGTGA 2365
```

RESULT 12
RNSTAT5BG

```
LOCUS       RNSTAT5BG               2671 bp    mRNA    linear   ROD 29-OCT-1997
DEFINITION  R.norvegicus mRNA for transcription factor Stat5b.
ACCESSION   X97541
VERSION     X97541.1  GI:1292887
KEYWORDS    prolactin-inducible protein; stat5b gene; transcription factor.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Luo,G. and Yu-Lee,L.
  TITLE     Transcriptional inhibition by Stat5: Differential activities at
            growth-related versus differentiation-specific promoters
  JOURNAL   J. Biol. Chem. 272, 26926-26934 (1997)
  MEDLINE   98001662
REFERENCE   2  (bases 1 to 2671)
  AUTHORS   Yu-Lee,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-1996) L. Yu-Lee, Baylor College of Medicine,
            Dept. of Medicine, Room 544E, One Baylor Plaza, Houston TX 77030,
            USA
FEATURES             Location/Qualifiers
     source          1. .2671
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Noble"
                     /db_xref="taxon:10116"
                     /clone="4"
                     /cell_line="Nb2 T cell"
                     /cell_type="T cell"
                     /tissue_type="lymph noce"
                     /clone_lib="Lambda ZAP phage"
     gene            1. .2671
                     /gene="stat5b"
     CDS             111. .2471
                     /gene="stat5b"
                     /codon_start=1
                     /protein_id="CAA66141.1"
                     /db_xref="GI:1292888"
                     /db_xref="GOA:P52632"
                     /db_xref="UniProt/Swiss-Prot:P52632"
                     /translation="MAMWIQAQQLQGDALHQMQALYGQHFPIEVRHYLSQWIESQAWD
                     SIDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQNTYDRCP
                     MELVRCIRHILYNEQRLVREANNGSSPAGSLADAMSQKHLQINQTFEELRLITQDTES
                     ELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLNPQERMSRETALQQKQVSLETWLQRE
                     AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
                     SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISALVTSTFII
                     EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQQAKSLLKNENTRNDY
                     SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKFTILFDSQFSV
                     GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
                     EALNMKFKAEVQSNRGLTKENLVFLAQKLFNSSSNHLEDYNSMSVSWSQFNRENLPGR
                     NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI
                     GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLNYLIYVFPDRPKDEVYSKY
                     YTPVPCEPATAKAADGYVKPQIKQVVPEFVNASADAGSGATYMDQAPSPVVCPQAHYN
                     MYPQNPDSVLDTDGDFDLEDTMDVARRVEELLGRPMDSQWIPHAQS"
ORIGIN
```

Query Match 84.2%; Score 1991.2; DB 10; Length 2671;
Best Local Similarity 90.4%; Pred. No. 0;
Matches 2138; Conservative 0; Mismatches 223; Indels 3; Gaps 1;



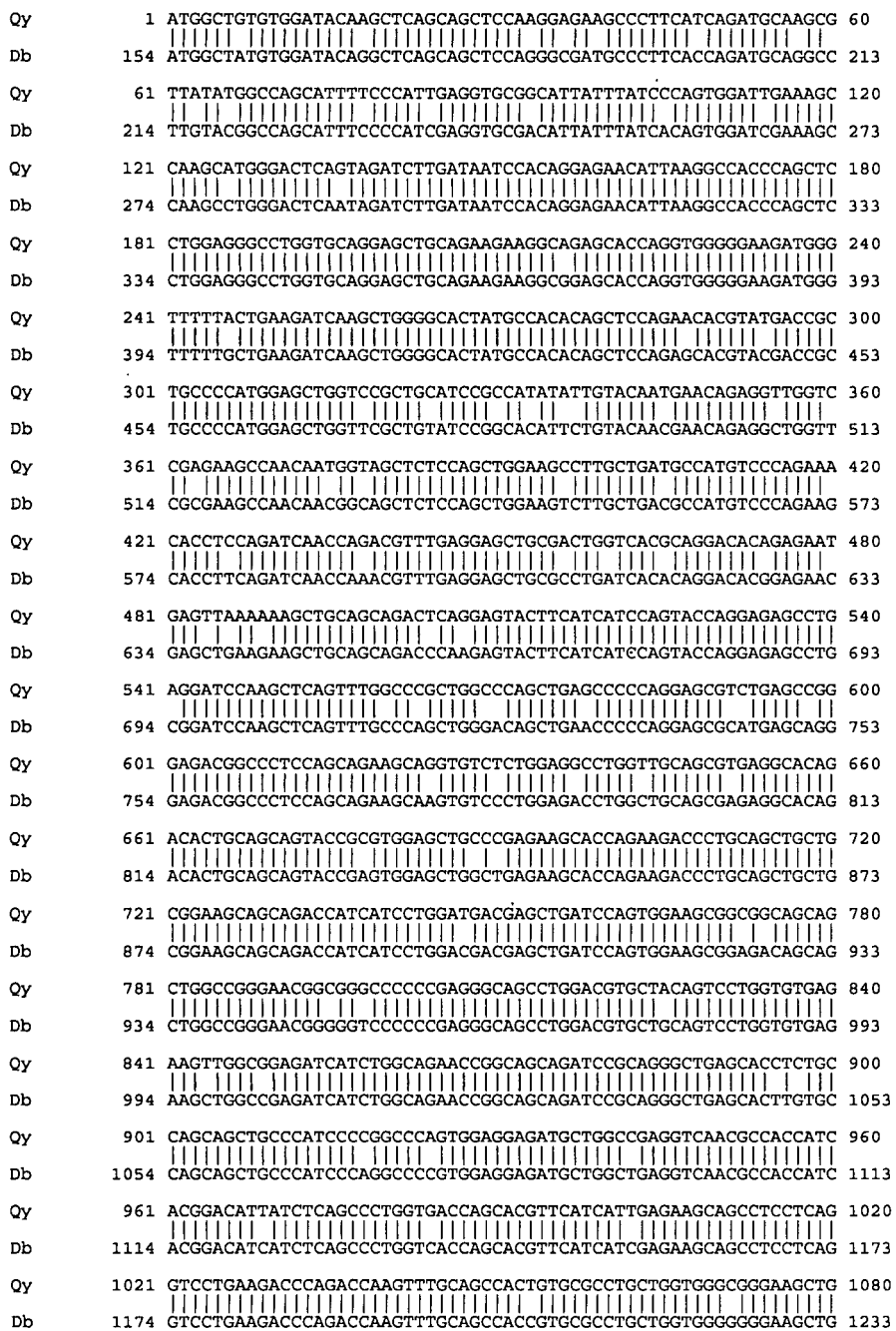


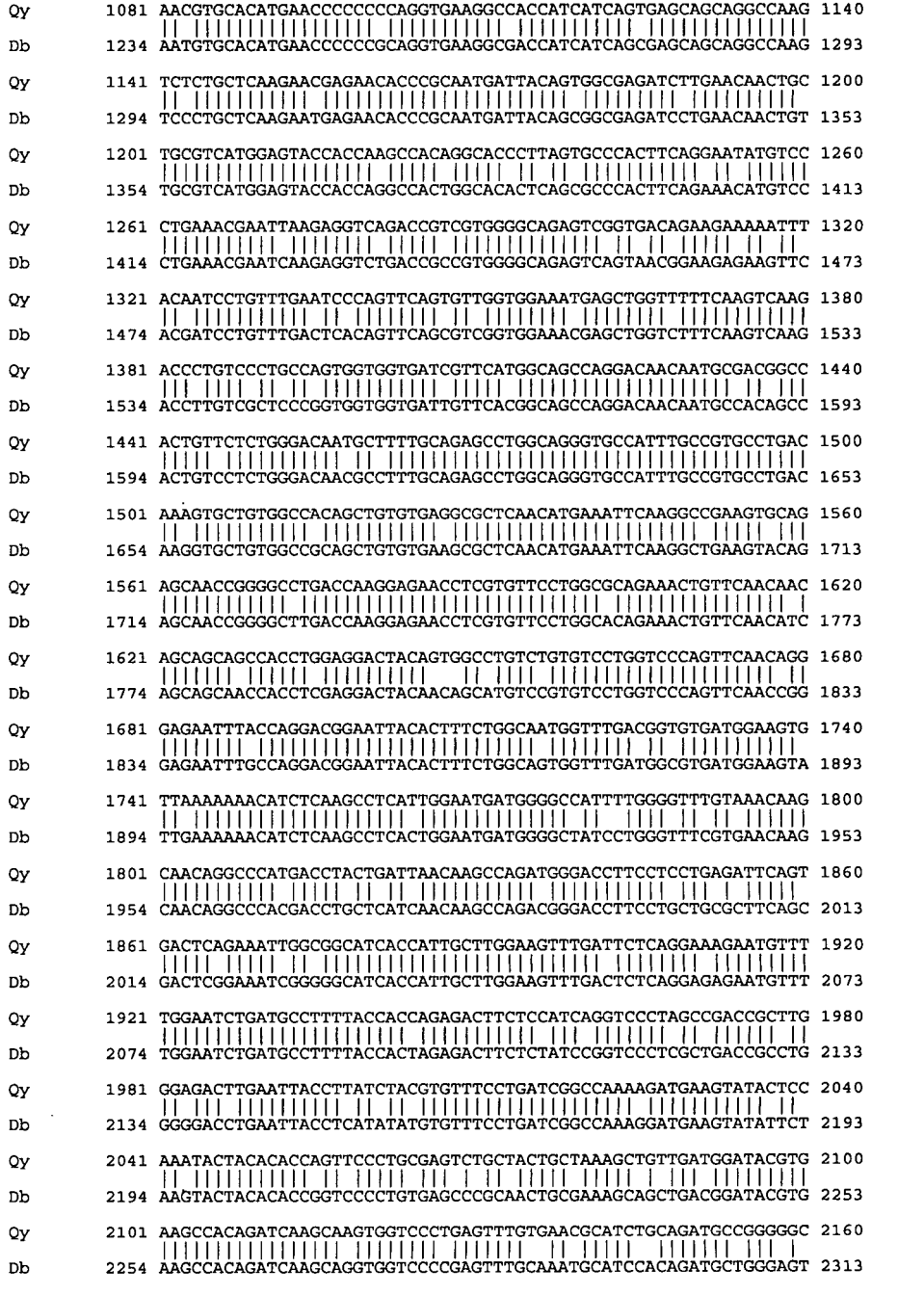
```
Db      2211 AAGCCGCAGATCAAGCAAGTGGTCCCCGAGTTTGTCAATGCTTCTGCAGATGCAGGGAGC 2270

Qy      2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
             |   ||||||| ||||||||||||||| || |||||||  ||||| || ||||||||||||
Db      2271 G---GCGCCACCTACATGGACCAGGCTCCTTCCCCAGTGGTGTGCCCTCAGGCTCACTAT 2327

Qy      2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
             |||||||||||||||| || ||||| ||||| || |||||||||| ||||||||||
Db      2328 AACATGTACCCACAGAATCCCGACTCTGTCCTCGATACCGATGGGGATTTCGATCTGGAA 2387

Qy      2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
             ||||| |||||||| || |||||||||| |||||| | |||||||| ||||||||||||
Db      2388 GACACGATGGACGTGGCACGGCGTGTGGAAGAGCTCTTAGGCCGGCCCATGGACAGTCAG 2447

Qy      2341 TGGATCCCGCACGCACAATCGTGA 2364
             |||||||| ||||| || || |||
Db      2448 TGGATCCCTCACGCGCAGTCATGA 2471
```

RESULT 13
BC024319

```
LOCUS       BC024319                4893 bp    mRNA    linear   ROD 06-OCT-2003
DEFINITION  Mus musculus signal transducer and activator of transcription 5B,
            mRNA (cDNA clone MGC:18421 IMAGE:4237132), complete cds.
ACCESSION   BC024319
VERSION     BC024319.1  GI:19354220
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4893)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 4893)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 25 Row: e Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 7242208.
FEATURES             Location/Qualifiers
     source          1. .4893
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:18421 IMAGE:4237132"
                     /tissue_type="Kidney, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Kid14"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .4893
                     /gene="Stat5b"
                     /db_xref="LocusID:20851"
                     /db_xref="MGI:103035"
     CDS             154. .2514
                     /gene="Stat5b"
                     /codon_start=1
                     /product="Stat5b protein"
                     /protein_id="AAH24319.1"
                     /db_xref="GI:19354221"
                     /db_xref="LocusID:20851"
                     /db_xref="MGI:103035"
                     /translation="MAMWIQAQQLQGDALHQMQALYGQHFPIEVRHYLSQWIESQAWD
                     SIDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQSTYDRCP
                     MELVRCIRHILYNEQRLVREANNGSSPAGSLADAMSQKHLQINQTFEELRLITQDTEN
                     ELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLNPQERMSRETALQQKQVSLETWLQRE
                     AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
                     SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISALVTSTFII
                     EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQQAKSLLKNENTRNDY
                     SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKFTILFDSQFSV
                     GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
                     EALNMKFKAEVQSNRGLTKENLVFLAQKLFNISSNHLEDYNSMSVSWSQFNRENLPGR
                     NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI
                     GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLNYLIYVFPDRPKDEVYSKY
                     YTPVPCEPATAKAADGYVKPQIKQVVPEFANASTDAGSGATYMDQAPSPVVCPQAHYN
                     MYPPNPDSVLDTDGDFDLEDTMDVARRVEELLGRPMDSQWIPHAQS"
     misc_feature    565. .1143
                     /gene="Stat5b"
                     /note="STAT; Region: STAT protein, all-alpha domain. STAT
                     proteins (Signal Transducers and Activators of
                     Transcription) are a family of transcription factors that
                     are specifically activated to regulate gene transcription
                     when cells encounter cytokines and growth factors. STAT
                     proteins also include an SH2 domain pfam00017"
                     /db_xref="CDD:pfam01017"
     misc_feature    1147. .1902
                     /gene="Stat5b"
                     /note="STAT_bind; Region: STAT protein, DNA binding
                     domain. STAT proteins (Signal Transducers and Activators
                     of Transcription) are a family of transcription factors
                     that are specifically activated to regulate gene
                     transcription when cells encounter cytokines and growth
                     factors. This family represents the DNA binding domain of
                     STAT, which has an ig-like fold. STAT proteins also
                     include an SH2 domain pfam00017"
                     /db_xref="CDD:pfam02864"
     misc_feature    1939. .2163
                     /gene="Stat5b"
                     /note="SH2; Region: SH2 domain"
                     /db_xref="CDD:pfam00017"
ORIGIN

  Query Match           83.8%;  Score 1981.6;  DB 10;  Length 4893;
  Best Local Similarity 90.2%;  Pred. No. 0;
```

Matches 2132; Conservative 0; Mismatches 229; Indels 3; Gaps 1;





```
Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
             |  ||||||| |||||||| ||||| || |||||||  ||||| || ||||||||||
Db       2314 G---GCGCCACCTACATGGATCAGGCTCCTTCCCCAGTCGTGTGCCCTCAGGCTCACTAC 2370

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
             ||||||||||||||  ||||| ||||| |||||||| |||||||||||||||||||||||||
Db       2371 AACATGTACCCACCCAACCCGGACTCCGTCCTTGATACCGATGGGGACTTCGATCTGGAA 2430

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
             ||||| |||||||| |||||||| ||||| |||||| | |||||||| ||||||||||||
Db       2431 GACACGATGGACGTGGCGCGGCGCGTGGAAGAGCTCTTAGGCCGGCCCATGGACAGTCAG 2490

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
             |||||||| |||||||| || |||
Db       2491 TGGATCCCTCACGCACAGTCATGA 2514
```

```
RESULT 14
AY040231
LOCUS       AY040231               2940 bp    mRNA    linear   ROD 15-MAR-2004
DEFINITION  Mus musculus signal transducer and activator of transcription 5B
            (Stat5b) mRNA, complete cds.
ACCESSION   AY040231
VERSION     AY040231.1  GI:21654824
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2940)
  AUTHORS   Davoodi-Semiromi,A., Laloraya,M., Kumar,G.P., Purohit,S., Jha,R.K.
            and She,J.-X.
  TITLE     A Mutant Stat5b with Weaker DNA Binding Affinity Defines a Key
            Defective Pathway in Nonobese Diabetic Mice
  JOURNAL   J. Biol. Chem. 279 (12), 11553-11561 (2004)
   PUBMED   14701862
REFERENCE   2  (bases 1 to 2940)
  AUTHORS   Davoodi-Semiromi,A., Laloraya,M., Kumar,P.G. and She,J.-X.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUN-2001) Pathology, University of Florida,
            Gainesville, FL 32610, USA
FEATURES             Location/Qualifiers
     source          1. .2940
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="NOD"
                     /db_xref="taxon:10090"
                     /chromosome="11"
     gene            1. .2940
                     /gene="Stat5b"
     CDS             484. .2844
                     /gene="Stat5b"
                     /codon_start=1
                     /product="signal transducer and activator of transcription
                     5B"
                     /protein_id="AAK74074.1"
                     /db_xref="GI:21654825"
                     /translation="MAMWIQAQQLQGDALHQMQALYGQHFPIEVRHYLSQWIESQAWD
                     SIDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQSTYDRCP
                     MELVRCIRHILYNEQRLVREANNGSSPAGSLADAMSQKHLQINQTFEELRLITQDTEN
                     ELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLNPQERMSRETALQQKQVSLETWLQRE
                     AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
                     SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISAMVTSTFII
                     EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQQAKSLLKNENTRNDY
                     SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKFTILFDSQFSV
                     GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
                     EALNMKFKAEVQSNRGLTKENLVFLAQKLFNISSNHLEDYNSMSVSWSQFNRENLPGR
                     NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI
                     GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLNYLIYVFPDRPKDEVYSKY
                     YTPVPCEPATAKAADGYVKPQIKQVVPEFANASTDAGSGATYMDQAPSPVVCPQAHYN
                     MYPPNPDSVLDTDGDFDLEDTMDVARRVEELLGRPMDSQWIPHAQS"
ORIGIN
```

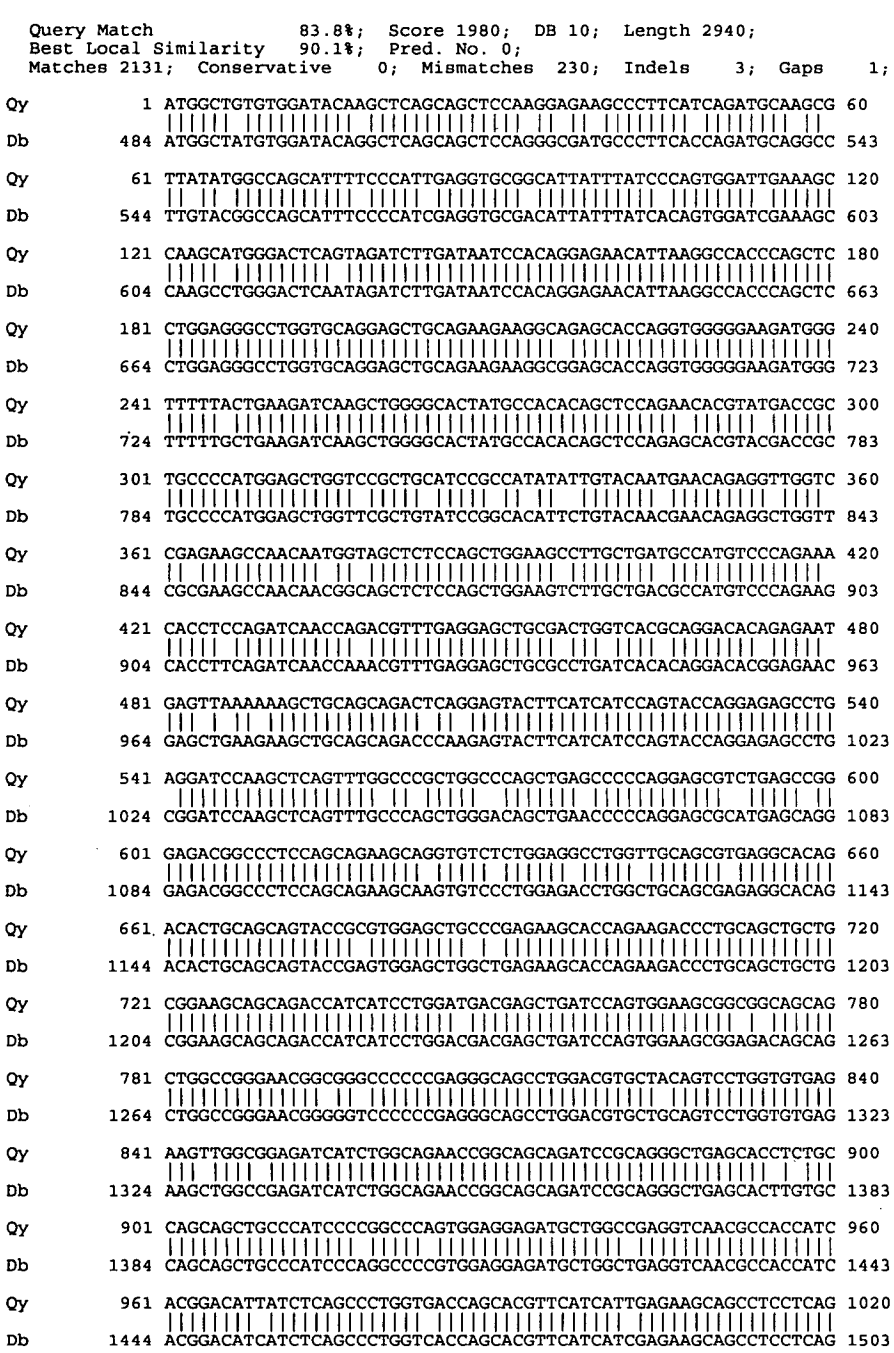
```
Query Match           83.8%;  Score 1980;  DB 10;  Length 2940;
Best Local Similarity 90.1%;  Pred. No. 0;
Matches 2131;  Conservative    0;  Mismatches  230;  Indels    3;  Gaps    1;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              |||||| |||||||||| |||||||||||||| || || |||||||| |||||||| ||
Db        484 ATGGCTATGTGGATACAGGCTCAGCAGCTCCAGGGCGATGCCCTTCACCAGATGCAGGCC 543

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              || || |||||||||||| ||||| |||||||| |||||||||||| |||||||| ||||||
Db        544 TTGTACGGCCAGCATTTCCCCATCGAGGTGCGACATTATTTATCACAGTGGATCGAAAGC 603

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||| ||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        604 CAAGCCTGGGACTCAATAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 663

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        664 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCGGAGCACCAGGTGGGGGAAGATGGG 723

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        724 TTTTTGCTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAGCACGTACGACCGC 783

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              |||||||||||||||| ||||| ||||| || ||  ||||||| ||||||||| ||||
Db        784 TGCCCCATGGAGCTGGTTCGCTGTATCCGGCACATTCTGTACAACGAACAGAGGCTGGTT 843

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              || |||||||||| || |||||||||||||||| |||||||| ||||||||||||||||
Db        844 CGCGAAGCCAACAACGGCAGCTCTCCAGCTGGAAGTCTTGCTGACGCCATGTCCCAGAAG 903

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||| |||||||||| |||||||||||||||| ||| |||| ||||||| |||||
Db        904 CACCTTCAGATCAACCAAACGTTTGAGGAGCTGCGCCTGATCACACAGGACACGGAGAAC 963

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||| | || |||||||||||||| || ||||||||||||||||||||||||||||||||
Db        964 GAGCTGAAGAAGCTGCAGCAGACCCAAGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 1023

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              |||||||||||||||| || ||||| ||||||| |||||||||||  ||||| ||
Db       1024 CGGATCCAAGCTCAGTTTGCCCAGCTGGGACAGCTGAACCCCCAGGAGCGCATGAGCAGG 1083

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              |||||||||||||||||||| ||||| |||||| ||||| |||||| ||||||||||
Db       1084 GAGACGGCCCTCCAGCAGAAGCAAGTGTCCCTGGAGACCTGGCTGCAGCGAGAGGCACAG 1143

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              |||||||||||||||| ||||||||| | ||||||||||||||||||||||||||||||
Db       1144 ACACTGCAGCAGTACCGAGTGGAGCTGGCTGAGAAGCACCAGAAGACCCTGCAGCTGCTG 1203

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              |||||||||||||||||||||||| |||||||||||||||||||||| | ||||||
Db       1204 CGGAAGCAGCAGACCATCATCCTGGACGACGAGCTGATCCAGTGGAAGCGGAGACAGCAG 1263

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||| |||||||||||||||||||||||||||||| ||||||||||||||
Db       1264 CTGGCCGGGAACGGGGGTCCCCCCGAGGGCAGCCTGGACGTGCTGCAGTCCTGGTGTGAG 1323

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||| |||| ||||||||||||||||||||||||||||||||||||||||||| | |||
Db       1324 AAGCTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACTTGTGC 1383

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              |||||||||||||||| ||||| |||||||||||||||| ||||||||||||||||||||
Db       1384 CAGCAGCTGCCCATCCCAGGCCCCGTGGAGGAGATGCTGGCTGAGGTCAACGCCACCATC 1443

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              |||||||| |||||||| |||| |||||||||||||||| ||||||||||||||||||
Db       1444 ACGGACATCATCTCAGCCATGGTCACCAGCACGTTCATCATCGAGAAGCAGCCTCCTCAG 1503

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
```

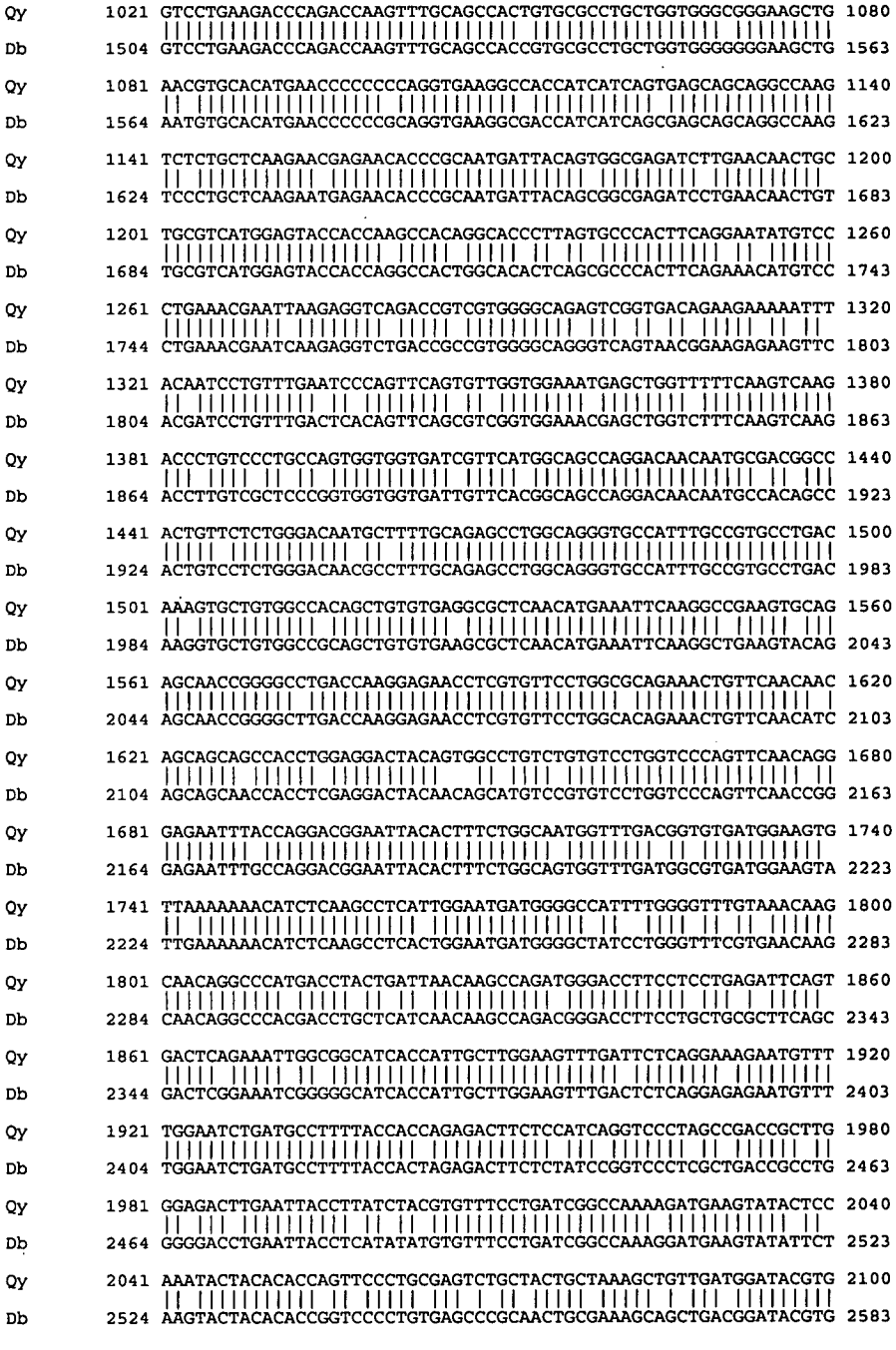
```
Db       1504 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACCGTGCGCCTGCTGGTGGGGGGGAAGCTG 1563

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140

Db       1564 AATGTGCACATGAACCCCCCGCAGGTGAAGGCGACCATCATCAGCGAGCAGCAGGCCAAG 1623

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200

Db       1624 TCCCTGCTCAAGAATGAGAACACCCGCAATGATTACAGCGGCGAGATCCTGAACAACTGT 1683

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260

Db       1684 TGCGTCATGGAGTACCACCAGGCCACTGGCACACTCAGCGCCCACTTCAGAAACATGTCC 1743

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320

Db       1744 CTGAAACGAATCAAGAGGTCTGACCGCCGTGGGGCAGAGTCAGTAACGGAAGAGAAGTTC 1803

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380

Db       1804 ACGATCCTGTTTGACTCACAGTTCAGCGTCGGTGGAAACGAGCTGGTCTTTCAAGTCAAG 1863

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440

Db       1864 ACCTTGTCGCTCCCGGTGGTGGTGATTGTTCACGGCAGCCAGGACAACAATGCCACAGCC 1923

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500

Db       1924 ACTGTCCTCTGGGACAACGCCTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1983

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560

Db       1984 AAGGTGCTGTGGCCGCAGCTGTGTGAAGCGCTCAACATGAAATTCAAGGCTGAAGTACAG 2043

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620

Db       2044 AGCAACCGGGGCTTGACCAAGGAGAACCTCGTGTTCCTGGCACAGAAACTGTTCAACATC 2103

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680

Db       2104 AGCAGCAACCACCTCGAGGACTACAACAGCATGTCCGTGTCCTGGTCCCAGTTCAACCGG 2163

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740

Db       2164 GAGAATTTGCCAGGACGGAATTACACTTTCTGGCAGTGGTTTGATGGCGTGATGGAAGTA 2223

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800

Db       2224 TTGAAAAAACATCTCAAGCCTCACTGGAATGATGGGGCTATCCTGGGTTTCGTGAACAAG 2283

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860

Db       2284 CAACAGGCCCACGACCTGCTCATCAACAAGCCAGACGGGACCTTCCTGCTGCGCTTCAGC 2343

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920

Db       2344 GACTCGGAAATCGGGGGCATCACCATTGCTTGGAAGTTTGACTCTCAGGAGAGAATGTTT 2403

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980

Db       2404 TGGAATCTGATGCCTTTTACCACTAGAGACTTCTCTATCCGGTCCCTCGCTGACCGCCTG 2463

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040

Db       2464 GGGGACCTGAATTACCTCATATATGTGTTTCCTGATCGGCCAAAGGATGAAGTATATTCT 2523

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100

Db       2524 AAGTACTACACACCGGTCCCCTGTGAGCCCGCAACTGCGAAAGCAGCTGACGGATACGTG 2583

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160

Db       2584 AAGCCACAGATCAAGCAGGTGGTCCCCGAGTTTGCAAATGCATCCACAGATGCTGGGAGT 2643

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220

Db       2644 G---GCGCCACCTACATGGATCAGGCTCCTTCCCCAGTCGTGTGCCCTCAGGCTCACTAC 2700

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280

Db       2701 AACATGTACCCACCCAACCCGGACTCCGTCCTTGATACCGATGGGGACTTCGATCTGGAA 2760

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340

Db       2761 GACACGATGGACGTGGCGCGGCGCGTGGAAGAGCTCTTAGGCCGGCCCATGGACAGTCAG 2820

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364

Db       2821 TGGATCCCTCACGCACAGTCATGA 2844
```

```
RESULT 15
AY044901
LOCUS       AY044901               2940 bp    mRNA    linear   ROD 01-JUL-2002
DEFINITION  Mus musculus strain BALB/c signal transducer and activator of
            transcription 5B (Stat5b) mRNA, complete cds.
ACCESSION   AY044901
VERSION     AY044901.1  GI:21654902
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2940)
  AUTHORS   Davoodi-Semiromi,A., Laloraya,M., Kumar,P. and She,J.-X.
  TITLE     Mutation detection in murine STAT gene family: identification and
            characterization of a novel point mutation in DNA binding domain of
            Stat5b in NOD
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2940)
  AUTHORS   Davoodi-Semiromi,A., Laloraya,M., Kumar,P. and She,J.-X.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2001) Pathology, University of Florida,
            Gainesville, FL 32610, USA
FEATURES             Location/Qualifiers
     source          1. .2940
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /chromosome="11"
     gene            1. .2940
                     /gene="Stat5b"
     CDS             484. .2844
                     /gene="Stat5b"
                     /note="transcription factor"
                     /codon_start=1
                     /product="signal transducer and activator of transcription
                     5B"
                     /protein_id="AAK97791.1"
                     /db_xref="GI:21654903"
                     /translation="MAMWIQAQQLQGDALHQMQALYGQHFPIEVRHYLSQWIESQAWD
                     SIDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQSTYDRCP
                     MELVRCIRHILYNEQRLVREANNGSSPAGSLADAMSQKHLQINQTFEELRLITQDTEN
                     ELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLNPQERMSRETALQQKQVSLETWLQRE
                     AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLAGNGGPPEGSLDVLQ
                     SWCEKLAEIIWQNRQQIRRAEHLCQQLPIPGPVEEMLAEVNATITDIISALVTSTFII
                     EKQPPQVLKTQTKFAATVRLLVGGKLNVHMNPPQVKATIISEQQAKSLLKNENTRNDY
                     SGEILNNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAGSVTEEKFTILFDSQFSV
                     GGNELVFQVKTLSLPVVVIVHGSQDNNATATVLWDNAFAEPGRVPFAVPDKVLWPQLC
                     EALNMKFKAEVQSNRGLTKENLVFLAQKLFNISSNHLEDYNSMSVSWSQFNRENLPGR
                     NYTFWQWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI
                     GGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADRLGDLNYLIYVFPDRPKDEVYSKY
                     YTPVPCEPATAKAADGYVKPQIKQVVPEFANASTDAGSGATYMDQAPSPVVCPQAHYN
                     MYPPNPDSVLDTDGDFDLEDTMDVARRVEELLGRPMDSQWIPHAQS"
```

ORIGIN

Query Match 83.8%; Score 1980; DB 10; Length 2940;
Best Local Similarity 90.1%; Pred. No. 0;
Matches 2131; Conservative 0; Mismatches 230; Indels 3; Gaps 1;





```
Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
             ||||||||||||||||| |||||||| |||||||  || |||||  ||||||| ||| |
Db       2584 AAGCCACAGATCAAGCAGGTGGTCCCCGAGTTTGCAAATGCATCCACAGATGCTGGGAGT 2643

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
             |   ||||||| |||||||| ||||| || |||||||  ||||| || ||||||||||||
Db       2644 G---GCGCCACCTACATGGATCAGGCTCCTTCCCCAGTCGTGTGCCCTCAGGCTCACTAC 2700

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
             |||||||||||||  ||||| ||||| |||||||| ||||||||||||||||||||||||
Db       2701 AACATGTACCCACCCAACCCGGACTCCGTCCTTGATACCGATGGGGACTTCGATCTGGAA 2760

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
             ||||| |||||||| |||||||| ||||| |||||| | |||||||| ||||||||||||
Db       2761 GACACGATGGACGTGGCGCGGCGCGTGGAAGAGCTCTTAGGCCGGCCCATGGACAGTCAG 2820

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
             |||||||| |||||||| || |||
Db       2821 TGGATCCCTCACGCACAGTCATGA 2844
```

Search completed: May 17, 2005, 06:14:57
Job time : 16237 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 23:15:04 ; Search time 11360 Seconds
(without alignments)
7921.121 Million cell updates/sec

Title: US-10-052-482-162
Perfect score: 2364
Sequence: 1 atggctgtgtggatacaagc.....tcccgcacgcacaatcgtga 2364

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1970.6	83.4	4986	3	AK045531	AK045531 Mus muscu
	2	956.8	40.5	1151	1	AL513644	AL513644 AL513644
	3	923.4	39.1	1084	4	BM449533	BM449533 AGENCOURT
c	4	862.8	36.5	1079	5	BX427445	BX427445 BX427445
	5	841	35.6	929	7	CO648803	CO648803 ILLUMIGEN
	6	820.6	34.7	1071	4	BM450408	BM450408 AGENCOURT
	7	809.2	34.2	877	5	BQ932442	BQ932442 AGENCOURT
	8	807.6	34.2	1025	5	BX401753	BX401753 BX401753
	9	805.8	34.1	909	5	BU508310	BU508310 AGENCOURT
	10	804	34.0	895	5	BQ706867	BQ706867 AGENCOURT
	11	780	33.0	879	5	BQ706430	BQ706430 AGENCOURT
	12	778.4	32.9	978	5	BX431739	BX431739 BX431739
	13	767	32.4	917	5	BQ941072	BQ941072 AGENCOURT
c	14	753	31.9	991	1	AL576279	AL576279 AL576279
	15	743.2	31.4	771	7	CF995633	CF995633 AGENCOURT
	16	734.6	31.1	742	5	BX499735	BX499735 DKFZp779P
	17	734.2	31.1	827	7	CO740375	CO740375 S1LL07a26
	18	724.6	30.7	975	1	AL553220	AL553220 AL553220
	19	717.4	30.3	755	5	BX114538	BX114538 BX114538
	20	700.6	29.6	983	5	BX427446	BX427446 BX427446
	21	700	29.6	886	5	BQ670478	BQ670478 AGENCOURT
	22	664.4	28.1	756	4	BG490357	BG490357 602519369
	23	664	28.1	985	5	BX422302	BX422302 BX422302
	24	658.2	27.8	730	5	BP147881	BP147881 BP147881
	25	647.4	27.4	758	7	CF535503	CF535503 UI-M-GH0-
	26	647.4	27.4	907	1	AL521374	AL521374 AL521374
	27	646.2	27.3	853	5	BU187856	BU187856 AGENCOURT
c	28	642.6	27.2	1019	5	BX385196	BX385196 BX385196
	29	638.6	27.0	715	4	BI520110	BI520110 603071247
	30	628.8	26.6	937	1	AL540450	AL540450 AL540450
	31	628.4	26.6	681	4	BG394677	BG394677 602456629
	32	623.8	26.4	699	5	BP145200	BP145200 BP145200
	33	623.4	26.4	922	5	BQ939396	BQ939396 AGENCOURT
	34	621.6	26.3	756	7	CF531926	CF531926 UI-M-GH0-
	35	621.6	26.3	1155	6	CD506834	CD506834 CDA81-E09
	36	617	26.1	793	2	BE545849	BE545849 601071753
	37	616.8	26.1	712	7	CN460794	CN460794 UI-M-HB0-
	38	614.6	26.0	706	7	CF742767	CF742767 UI-M-HB0-
	39	606.4	25.7	617	7	CF137996	CF137996 UI-HF-BN0
	40	605.8	25.6	769	4	BI415585	BI415585 602988371
	41	604.4	25.6	720	4	BI650503	BI650503 603296865
	42	604.2	25.6	828	6	CD616432	CD616432 56075624J
	43	602	25.5	789	7	CK482172	CK482172 AGENCOURT
	44	589.4	24.9	811	2	BE253625	BE253625 601109070
	45	588.4	24.9	604	7	CN357151	CN357151 170005316

ALIGNMENTS

RESULT 1
AK045531
LOCUS AK045531 4986 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230209B11 product:signal transducer and activator of transcription 5B, full insert sequence.
ACCESSION AK045531
VERSION AK045531.1 GI:26090947
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 4986)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.

FEATURES Location/Qualifiers
source 1. .4986
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:B230209B11"
/db_xref="taxon:10090"
/clone="B230209B11"
/sex="male"
/tissue_type="corpora quadrigemina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature 340. .2701
/note="putative signal transducer and activator of transcription 5B (MGD|MGI:103035, GB|NM_011489, evidence: BLASTN, 99%, match=2574)"

ORIGIN

Query Match 83.4%; Score 1970.6; DB 3; Length 4986;
Best Local Similarity 90.1%; Pred. No. 0;
Matches 2132; Conservative 0; Mismatches 229; Indels 4; Gaps 2;

```
Db       1660 CACGATCCTGTTTGACTCACAGTTCAGCGTCGGTGGAAACGAGCTGGTCTTTCAAGTCAA 1719

Qy       1380 GACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGC 1439

Db       1720 GACCTTGTCGCTCCCGGTGGTGGTGATTGTTCACGGCAGCCAGGACAACAATGCCACAGC 1779

Qy       1440 CACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGA 1499

Db       1780 CACTGTCCTCTGGGACAACGCCTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGA 1839

Qy       1500 CAAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCA 1559

Db       1840 CAAGGTGCTGTGGCCGCAGCTGTGTGAAGCGCTCAACATGAAATTCAAGGCTGAAGTACA 1899

Qy       1560 GAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAA 1619

Db       1900 GAGCAACCGGGGCTTGACCAAGGAGAACCTCGTGTTCCTGGCACAGAAACTGTTCAACAT 1959

Qy       1620 CAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAG 1679

Db       1960 CAGCAGCAACCACCTCGAGGACTACAACAGCATGTCCGTGTCCTGGTCCCAGTTCAACCG 2019

Qy       1680 GGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGT 1739

Db       2020 GGAGAATTTGCCAGGACGGAATTACACTTTCTGGCAGTGGTTTGATGGCGTGATGGAAGT 2079

Qy       1740 GTTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAA 1799

Db       2080 ATTGAAAAAACATCTCAAGCCTCACTGGAATGATGGGGCTATCCTGGGTTTCGTGAACAA 2139

Qy       1800 GCAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAG 1859

Db       2140 GCAACAGGCCCACGACCTGCTCATCAACAAGCCAGACGGGACCTTCCTGCTGCGCTTCAG 2199

Qy       1860 TGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTT 1919

Db       2200 CGACTCGGAAATCGGGGGCATCACCATTGCTTGGAAGTTTGACTCTCAGGAGAGAATGTT 2259

Qy       1920 TTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTT 1979

Db       2260 TTGGAATCTGATGCCTTTTACCACTAGAGACTTCTCTATCCGGTCCCTCGCTGACCGCCT 2319

Qy       1980 GGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTC 2039

Db       2320 GGGGGACCTGAATTACCTCATATATGTGTTTCCTGATCGGCCAAAGGATGAAGTATATTC 2379

Qy       2040 CAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGT 2099

Db       2380 TAAGTACTACACACCGGTCCCCTGTGAGCCCGCAACTGCGAAAGCAGCTGACGGATACGT 2439

Qy       2100 GAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGG 2159

Db       2440 GAAGCCACAGATCAAGCAGGTGGTCCCCGAGTTTGCAAATGCATCCACAGATGCTGGGAG 2499

Qy       2160 CGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTA 2219

Db       2500 TG---GCGCCACCTACATGGATCAGGCTCCTTCCCCAGTCGTGTGCCCTCAGGCTCACTA 2556

Qy       2220 TAACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGA 2279

Db       2557 CAACATGTACCCACCCAACCCGGACTCCGTCCTTGATACCGATGGGGACTTCGATCTGGA 2616

Qy       2280 GGACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCA 2339

Db       2617 AGACACGATGGACGTGGCGCGGCGCGTGGAAGAGCTCTTAGGCCGGCCCATGGACAGTCA 2676

Qy       2340 GTGGATCCCGCACGCACAATCGTGA 2364

Db       2677 GTGGATCCCTCACGCACAGTCATGA 2701


RESULT 2
AL513644
```

```
LOCUS       AL513644               1151 bp    mRNA    linear   EST 23-MAR-2004
DEFINITION  AL513644 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            XCL0BB001ZG12 5-PRIME, mRNA sequence.
ACCESSION   AL513644
VERSION     AL513644.3  GI:45650618
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutèleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1151)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31037647.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 8244.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=XCL0BB001ZG12RP1&c=8244.r.
FEATURES             Location/Qualifiers
     source          1. .1151
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="XCL0BB001ZG12"
                     /tissue_type="NEUROBLASTOMA"
                     /clone_lib="Homo sapiens NEUROBLASTOMA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN
```

```
  Query Match           40.5%;  Score 956.8;  DB 1;  Length 1151;
  Best Local Similarity   99.5%;  Pred. No. 2e-227;
  Matches  969;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       1391 TGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCCACTGTTCTCT 1450

Db          1 TGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCCACTGTTCTCT 60

Qy       1451 GGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGCTGT 1510

Db         61 GGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGCTGT 120

Qy       1511 GGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACCGGG 1570

Db        121 GGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACCGGG 180

Qy       1571 GCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCAGCC 1630

Db        181 GCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCAGCC 240

Qy       1631 ACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGGAGAATTTAC 1690

Db        241 ACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGGAGAATTTAC 300

Qy       1691 CAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTGTTAAAAAAAC 1750

Db        301 CAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTGTTAAAAAAAC 360

Qy       1751 ATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGCAACAGGCCC 1810

Db        361 ATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGCAACAGGCCC 420
```

```
Qy       1811 ATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGTGACTCAGAAA 1870
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGTGACTCAGAAA 480

Qy       1871 TTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTTGGAATCTGA 1930
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTTGGAATCTGA 540

Qy       1931 TGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTGGGAGACTTGA 1990
             |||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        541 TGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCGACCGCTTGGGAGACTTGA 600

Qy       1991 ATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCCAAATACTACA 2050
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCCAAATACTACA 660

Qy       2051 CACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTGAAGCCACAGA 2110
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTGAAGCCACAGA 720

Qy       2111 TCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGCGGCAGCGCCA 2170
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        721 TCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGCGGCA-CGCCA 779

Qy       2171 CGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTATAACATGTACC 2230
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        780 CGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTYCCCAGGCTCACTATAACATGTACC 839

Qy       2231 CACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAGGACACAATGG 2290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAGGACACAATGG 899

Qy       2291 ACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAGTGGATCCCGC 2350
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 ACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAGTGGATCCCGC 959

Qy       2351 ACGCACAATCGTGA 2364
             ||||||||||||||
Db        960 ACGCACAATCGTGA 973
```

RESULT 3
BM449533

```
LOCUS       BM449533               1084 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6386129 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5526703
            5', mRNA sequence.
ACCESSION   BM449533
VERSION     BM449533.1  GI:18498573
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1084)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12200  row: m  column: 08
            High quality sequence stop: 719.
FEATURES             Location/Qualifiers
     source          1. .1084
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5526703"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN
```

```
  Query Match             39.1%;  Score 923.4;  DB 4;  Length 1084;
  Best Local Similarity   98.0%;  Pred. No. 4.4e-219;
  Matches  956;  Conservative    0;  Mismatches   17;  Indels    3;  Gaps    2;

Qy       1010 AGCCTCCTCAGGTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGG 1069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCTCCTCAGGTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGG 60

Qy       1070 GCGGGAAGCTGAACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGC 1129
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGGGAAGCTGAACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGC 120

Qy       1130 AGCAGGCCAAGTCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCT 1189
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AGCAGGCCAAGTCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCT 180

Qy       1190 TGAACAACTGCTGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCA 1249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TGAACAACTGCTGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCA 240

Qy       1250 GGAATATGTCCCTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAG 1309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GGAATATGTCCCTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAG 300

Qy       1310 AAGAAAAATTTACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTT 1369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AAGAAAAATTTACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTT 360

Qy       1370 TTCAAGTCAAGACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACA 1429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TTCAAGTCAAGACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACA 420

Qy       1430 ATGCGACGGCCACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTG 1489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 ATGCGACGGCCACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTG 480

Qy       1490 CCGTGCCTGACAAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGG 1549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CCGTGCCTGACAAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGG 540

Qy       1550 CCGAAGTGCAGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAAC 1609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CCGAAGTGCAGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAAC 600

Qy       1610 TGTTCAACAACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCC 1669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 TGTTCAACAACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCC 660

Qy       1670 AGTTCAACAGGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTG 1729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AGTTCAACAGGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTG 720

Qy       1730 TGATGGAAGTGTTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCA--TTTTGGG 1787
             |||||| |||||||||||||||||||||||||||||||||||||| ||||||| |||| ||| |||
Db         721 TGATGGGAGTGTTAAAAAAACATCTCAAGCCTCATTGGGATGATGGNGCCACTTTTGGGG 780

Qy       1788 GTTTGTAAACAAGCAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCT 1847
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         781 GTTTGTAAACAAGCAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCT 840

Qy       1848 CCTGAGATTCAGTGACTCAGAAATTGGCGGCATCACCATTGCTT-GGAAGTTTGATTCTC 1906
             |||||||||||||||||||||||||||||||||||| || |||||| ||||| |||||||||
```

```
Db        841 CCTGAGATTCAGTGACTCAGAAATTGGCGGCATCCCCCTTGCTTGGGAAGATTGATTCTC 900

Qy       1907 AGGAAAGAATGTTTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCC 1966
              ||||||||| |||||||||| |||||||||||||||| | ||||||||| ||||||
Db        901 AGGAAAGAAGTTTTTGGAATCGGATGCCTTTTACCACCAAAAACTTCTCCATTCGGTCCC 960

Qy       1967 TAGCCGACCGCTTGGG 1982
              || ||||||| ||||||
Db        961 TAACCGACCCCTTGGG 976


RESULT 4
BX427445/c
LOCUS       BX427445                1079 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX427445 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
            CS0DM010YO20 3-PRIME, mRNA sequence.
ACCESSION   BX427445
VERSION     BX427445.2  GI:47011175
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30782650.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 8244.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0AM010BH10NP1&c=8244.r.
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM010YO20"
                     /tissue_type="FETAL LIVER"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL LIVER"
                     /note="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             36.5%;  Score 862.8;  DB 5;  Length 1079;
  Best Local Similarity   95.4%;  Pred. No. 5.9e-204;
  Matches  883;  Conservative   22;  Mismatches   19;  Indels    2;  Gaps    2;

Qy       1439 CCACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTG 1498
              |: |  || |:|  |   |  ::||||||||:| :|||: ||||||||||||||:||||
Db       1030 CSGCCATTTTYTYTGGACAATSYTTTGCAGASCTKGGCR-GGTGCCATTTGCCGTSCCTG 972

Qy       1499 ACAAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGC 1558
              |||||||| ::||||||||||||:|||:|||||||||:||||||||||:|||||||||
Db        971 ACAAAGTG-YKTGGCCACAGCTGTKTGARGCGCTCAACWTGAAATTCAAGSCCGAAGTGC 913

Qy       1559 AGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACA 1618
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        912 AGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTYCCTGGCGCAGAAACTGTTCAACA 853

Qy       1619 ACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACA 1678
              |||||||||||||||||||:||||:  |: |:|||:|||||:||||||||||||||||
Db        852 ACAGCAGCAGCCACCTGGAGRACTAYTTTKTCYTGTYTGTGTYCTGGTCCCAGTTCAACA 793

Qy       1679 GGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAG 1738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 GGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAG 733

Qy       1739 TGTTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACA 1798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 TGTTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACA 673

Qy       1799 AGCAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCA 1858
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        672 AGCAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCA 613

Qy       1859 GTGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGT 1918
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GTCACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGT 553

Qy       1919 TTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCT 1978
              |||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db        552 TTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCGACCGCT 493

Qy       1979 TGGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACT 2038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TGGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACT 433

Qy       2039 CCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACG 2098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACG 373

Qy       2099 TGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGG 2158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGG 313

Qy       2159 GCGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACT 2218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GCGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACT 253

Qy       2219 ATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGG 2278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 ATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGG 193

Qy       2279 AGGACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTC 2338
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        192 AGGACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGSACAGTC 133

Qy       2339 AGTGGATCCCGCACGCACAATCGTGA 2364
              |||||||||||||||||||||||||
Db        132 AGTGGATCCCGCACGCACAATCGTGA 107


RESULT 5
CO648803
LOCUS       CO648803                929 bp    mRNA    linear   EST 23-JUL-2004
DEFINITION  ILLUMIGEN_MCQ_42294 Katze_MMPB2 Macaca mulatta cDNA clone
            IBIUW:22257 5' similar to Bases 5 to 929 highly similar to human
            STAT5B (Hs.434992), mRNA sequence.
ACCESSION   CO648803
VERSION     CO648803.1  GI:50570297
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
  TITLE     Large-scale Rhesus Macaque cDNA Sequencing
  JOURNAL   Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
```

```
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.06.03. 715 Q20 bases. Library Preparation: Prof.
            Michael Katze Lab at University of Washington DNA Sequencing:
            Illumigen Biosciences Inc. For further information, see
            http://www.macaque.org
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 929   Std Error: 0.00
            Plate: CL000311  row: D  column: 02
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .929
                      /organism="Macaca mulatta"
                      /mol_type="mRNA"
                      /strain="Indian"
                      /db_xref="taxon:9544"
                      /clone="IBIUW:22257"
                      /sex="male"
                      /tissue_type="blood"
                      /cell_type="PBMC"
                      /dev_stage="adult"
                      /lab_host="Electromax DH10B"
                      /clone_lib="Katze_MMPB2"
                      /note="Vector: pDONR 222; Site_1: BsrG I; Site_2: BsrG I;
                      Created from CloneMiner cDNA Library Construction kit
                      (catalog #18249-029)"
ORIGIN

  Query Match             35.6%;  Score 841;  DB 7;  Length 929;
  Best Local Similarity   96.3%;  Pred. No. 1.6e-198;
  Matches  893;  Conservative    0;  Mismatches   30;  Indels    4;  Gaps    3;

Qy       1268 GAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTTACAATCC 1327
             | ||||||||  ||||||||||||||||||||||| ||||||||||||||||||||||||
Db          3 GGATTAAGAGATCAGACCGTCGTGGGGCAGAGTCAGTGACAGAAGAAAAATTTACAATCC 62

Qy       1328 TGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAGACCCTGT 1387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAGACCCTGT 122

Qy       1388 CCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCCACTGTTC 1447
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||| ||||
Db        123 CCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCAACGGCCACCGTTC 182

Qy       1448 TCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGC 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 TCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGC 242

Qy       1508 TGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACC 1567
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACC 302

Qy       1568 GGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCA 1627
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        303 GGGGCCTGACCAAGGAGAACCTTGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCA 362

Qy       1628 GCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGGAGAATT 1687
             |||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||
Db        363 GCCACCTGGAGGACTACAGCGGCCTGTCCGTGTCCTGGTCCCAGTTCAACAGGGAGAATT 422

Qy       1688 TACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTGTTAAAAA 1747
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        423 TACCAGGACGGAATTACACTTTCTGGCAGTGGTTTGACGGTGTGATGGAAGTGTTAAAAA 482

Qy       1748 AACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGCAACAGG 1807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGCAACAGG 542

Qy       1808 CCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGTGACTCAG 1867
             ||||||||||||| |||||||||||||||||||||||| |||||||||||||||||||
Db        543 CCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTGCTGAGATTCAGTGACTCAG 602

Qy       1868 AAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTTGGAATC 1927
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        603 AAATTGGCGGCATCACCATTGTTTGGAAGTTTGATTCTCAGGAAAGAATGTTTTGGAATC 662

Qy       1928 TGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTGGGAGACT 1987
             |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db        663 TGATGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCGACCGCTTGGGAGACT 722

Qy       1988 TGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCCAAATACT 2047
             |||||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db        723 TGAATTACCTTATCTACGTGTTTCCTGATCGGGCGAAAGATGAAGTATACTCCAAATACT 782

Qy       2048 ACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAG-CTGTTGATGGATACGTGAAGCCA 2106
             |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||||
Db        783 ACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCCCGTTGATGGATACGTGAAGCCA 842

Qy       2107 CAGATCAAGCAAGTGG-TCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGG--GGCGGC 2163
             || |||||||||||||| ||||||| ||||||||| | |||||| ||| ||||  |  |||
Db        843 CAAATCAAGCAAGTGGTTCCCTGAATTTGTGAACCCTTCTGCAAATGGCGGGGAGCGGGC 902

Qy       2164 AGCGCCACGTACATGGACCAGGCCCCC 2190
             |||||| | ||| ||||||| | ||||
Db        903 AGCGCCCCCTACCTGGACCAAGGCCCC 929
```

```
RESULT 6
BM450408
LOCUS       BM450408               1071 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6394143 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5494762
            5', mRNA sequence.
ACCESSION   BM450408
VERSION     BM450408.1  GI:18499448
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1071)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12120  row: j  column: 11
            High quality sequence stop: 633.
FEATURES              Location/Qualifiers
     source           1. .1071
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5494762"
                      /tissue_type="retinoblastoma"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_67"
                      /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.75 kb.   Library constructed by Life
                      Technologies."
ORIGIN

  Query Match             34.7%;  Score 820.6;  DB 4;  Length 1071;
  Best Local Similarity   96.2%;  Pred. No. 2e-193;
  Matches  862;  Conservative    0;  Mismatches   30;  Indels    4;  Gaps    2;
```

```
Qy       1284 CCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTTACAATCCTGTTTGAATCCCAGTT 1343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTTACAATCCTGTTTGAATCCCAGTT 60

Qy       1344 CAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAGACCCTGTCCCTGCCAGTGGTGGT 1403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAGACCCTGTCCCTGCCAGTGGTGGT 120

Qy       1404 GATCGTTCATGGCAGCCAGGACAACAATGCGACGGCCACTGTTCTCTGGGACAATGCTTT 1463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATCGTTCATGGCAGCCAGGACAACAATGCGACGGCCACTGTTCTCTGGGACAATGCTTT 180

Qy       1464 TGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGCTGTGGCCACAGCTGTG 1523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGCTGTGGCCACAGCTGTG 240

Qy       1524 TGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACCGGGGCCTGACCAAGGA 1583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACCGGGGCCTGACCAAGGA 300

Qy       1584 GAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCAGCCACCTGGAGGACTA 1643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCAGCCACCTGGAGGACTA 360

Qy       1644 CAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGGAGAATTTACCAGGACGGAATTA 1703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGGAGAATTTACCAGGACGGAATTA 420

Qy       1704 CACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTGTTAAAAAAACATCTCAAGCCTCA 1763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTGTTAAAAAAACATCTCAAGCCTCA 480

Qy       1764 TTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGCAACAGGCCCATGACCTACTGAT 1823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        481 TTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGCAACAGGCCCATGACCTACTCAT 540

Qy       1824 TAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGTGACTCAGAAATTGGCGGCATCAC 1883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGTGACTCAGAAATTGGCGGCATCAC 600

Qy       1884 CATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTTGGAATCTGATGCCTTTTACCAC 1943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTTGGAATCTGATGCCTTTTACCAC 660

Qy       1944 CAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTGGGAGACTTGAATTACCTTATCTA 2003
             |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGAGACTTCTCCATTCGGTCCCTAGCCGACCGCTTGGGAGACTTGAATTACCTTATCTA 720

Qy       2004 CGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCCAAATACTACACACCAG-TTCCCT 2062
             |||||||||||||||||  |||||||||||||||||||||||||||||||||| ||||||
Db        721 CGTGTTTCCTGATCGGCCANAAGATGAAGTATACTCCAAATACTACACACCAGTTTCCCT 780

Qy       2063 GCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTGAAGCCACAGATCAAGCAAGTGG 2122
             |||||||||||||||||||||||||||||||||| ||||||||||   |||||||||
Db        781 GCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACCTGAAGCCACAGAATCAGCAAGTGG 840

Qy       2123 TCCC---TGAGTTTGTGAACGCATCTGCAGATGCCGGGGGCGGCAGCGCCACGTAC 2175
              |||     | ||  ||| ||||||||  |   |  ||| | | ||  | |  ||
Db        841 GCCCCTGAAATTTGTTGACCGCATCTGCCAAATGCCCGGGGGCCGGCAACCCCCAC 896
```

```
RESULT 7
BQ932442
LOCUS       BQ932442                877 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8854276 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6473729
            5', mRNA sequence.
ACCESSION   BQ932442
VERSION     BQ932442.1  GI:22347796
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 877)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14009  row: 1  column: 18
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .877
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6473729"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN
```

```
  Query Match             34.2%;  Score 809.2;  DB 5;  Length 877;
  Best Local Similarity   98.9%;  Pred. No. 1.4e-190;
  Matches  825;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 60

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 120

Qy        481 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 180

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 240

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 300

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        301 ACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 360

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 420

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 480

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGTTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 540

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 600
```

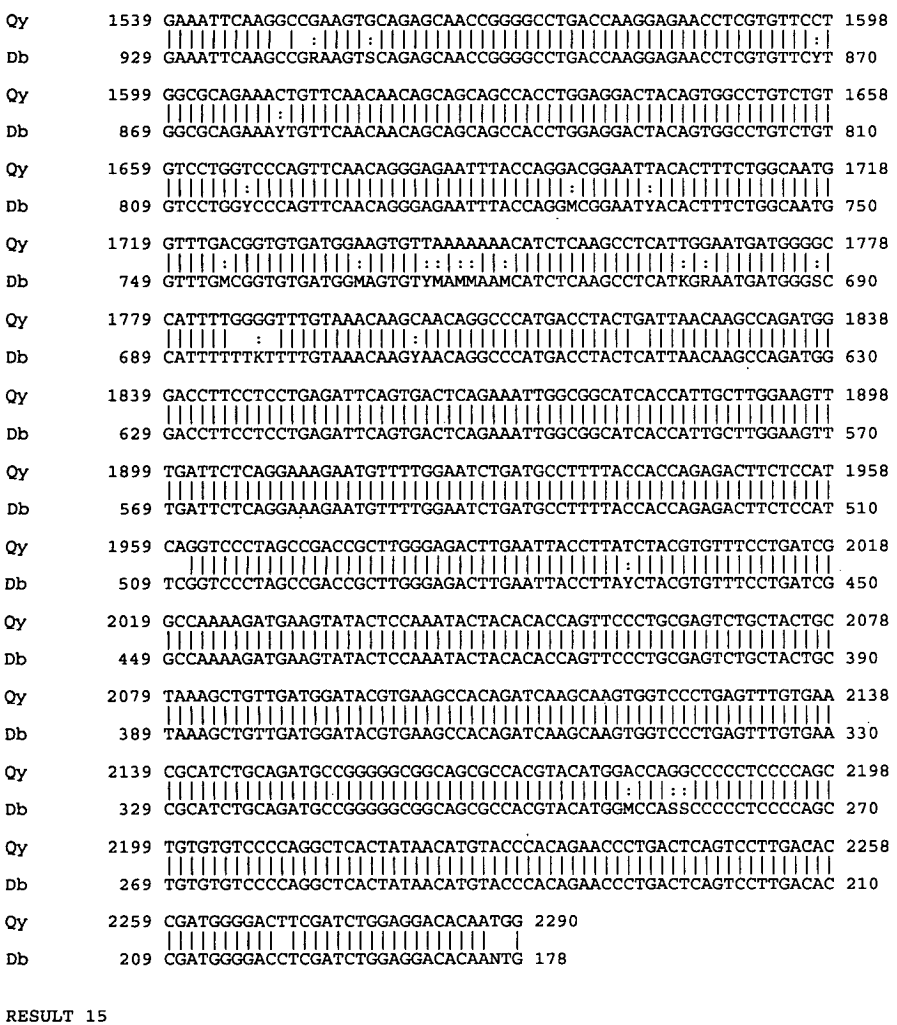
```
Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 660

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        661 GTCCTGAAGACCCAGACCAAGCTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 720

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              |||||| |||||||||||||||| |||||||||||||||||||||||||||||||| ||
Db        721 AACGTGGACATGAACCCCCCCCCAGGCGAAGGCCACCATCATCAGTGAGCAGCAGGCCCAG 780

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGT-GGCGAGATCTTGAA 1193
              |||||||||||||||||||| |||||||||| ||||| |||||||||||||||
Db        781 TCTCTGCTCAAGAACGAGAACCCCCGCAATGATAACAGTGGGCGAGATCTTGAA 834


RESULT 8
BX401753
LOCUS       BX401753               1025 bp    mRNA    linear   EST 28-APR-2004
DEFINITION  BX401753 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
            Homo sapiens cDNA clone CS0DL008YC13 5-PRIME, mRNA sequence.
ACCESSION   BX401753
VERSION     BX401753.2  GI:46835520
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1025)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30610754.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            8244.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DL008AB07QP1&c=8244.r.
FEATURES             Location/Qualifiers
     source          1. .1025
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DL008YC13"
                     /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
                     25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match           34.2%;  Score 807.6;  DB 5;  Length 1025;
  Best Local Similarity   97.2%;  Pred. No. 3.5e-190;
  Matches  846;  Conservative    6;  Mismatches   15;  Indels    3;  Gaps    3;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 207

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        208 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 267

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 327

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 387

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 447

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 507

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 567

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 627

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 687

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 747

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 807

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        808 ACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 867

Qy        721 CGGAAGCAGCAGACCATCATCCT-GGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCA 779
              ||||||::: |  | |||||||| ||||||||||||||||||||||||||||||||| ||
Db        868 CGGAAGMMRGAACCAATCATCCTKGGATGACGAGCTGATCCAGTGGAAGCGGCGGCA-CA 926

Qy        780 GCTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGA 839
              ||||| | || |  |||||||||||||||||||||||||||||||||||||||||||||
Db        927 GCTGGGCCGGGAACGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGA 986

Qy        840 GAAGTTGGCGGAGATCATCTGGCAGAACCG 869
              | ||||||  |||||||:|||:: | | |
Db        987 G-AGTTGGGCGAGATCATYTGGMRAACCGG 1015


RESULT 9
BU508310
LOCUS       BU508310               909 bp    mRNA    linear   EST 12-SEP-2002
DEFINITION  AGENCOURT_10094472 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502375
            5', mRNA sequence.
ACCESSION   BU508310
VERSION     BU508310.1  GI:22814543
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14058  row: f  column: 08
            High quality sequence stop: 638.
FEATURES             Location/Qualifiers
     source          1. .909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6502375"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match            34.1%;  Score 805.8;  DB 5;  Length 909;
  Best Local Similarity  97.9%;  Pred. No. 9.7e-190;
  Matches  890;  Conservative    0;  Mismatches   12;  Indels    7;  Gaps    7;

Qy       1048 GCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTGAACGTGCACATGAACCCCCCCCAGGTG 1107
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db             1 GCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTGAACGTGCACATGAACCCCCCACAGGTG 60

Qy       1108 AAGGCCACCATCATCAGTGAGCAGCAGGCCAAGTCTCTGCTCAAGAACGAGAACACCCGC 1167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 AAGGCCACCATCATCAGTGAGCAGCAGGCCAAGTCTCTGCTCAAGAACGAGAACACCCGC 120

Qy       1168 AATGATTACAGTGGCGAGATCTTGAACAACTGCTGCGTCATGGAGTACCACCAAGCCACA 1227
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGATTACAGTGGCGAGATCTTGAACAACTGCTGCGTCATGGAGTACCACCAAGCCACA 180

Qy       1228 GGCACCCTTAGTGCCCACTTCAGGAATATGTCCCTGAAACGAATTAAGAGGTCAGACCGT 1287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCACCCTTAGTGCCCACTTCAGGAATATGTCCCTGAAACGAATTAAGAGGTCAGACCGT 240

Qy       1288 CGTGGGGCAGAGTCGGTGACAGAAGAAAAATTTACAATCCTGTTTGAATCCCAGTTCAGT 1347
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTGGGGCAGAGTCGGTGACAGAAGAAAAATTTACAATCCTGTTTGAATCCCAGTTCAGT 300

Qy       1348 GTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAGACCCTGTCCCTGCCAGTGGTGGTGATC 1407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAGACCCTGTCCCTGCCAGTGGTGGTGATC 360

Qy       1408 GTTCATGGCAGCCAGGACAACAATGCGACGGCCACTGTTCTCTGGGACAATGCTTTTGCA 1467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTCATGGCAGCCAGGACAACAATGCGACGGCCACTGTTCTCTGGGACAATGCTTTTGCA 420

Qy       1468 GAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGCTGTGGCCACAGCTGTGTGAG 1527
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGCTGTGGCCACAGCTGTGTGAG 480

Qy       1528 GCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACCGGGGCCTGACCAAGGAGAAC 1587
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACCGGGGCCTGACCAAGGAGAAC 540

Qy       1588 CTCGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCAGCCACCTGGAGGACTACAGT 1647
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCAGCCACCTGGAGGACTACAGT 600

Qy       1648 GGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGGAGAATTTACCAGGACGGAATTACACT 1707
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGGAGAATTTACCAGGACGGAATTACACT 660

Qy       1708 TTCTGGCAATGGTTTGACGGTGTGATGGAAGTGTTAAAAAAACATCTCAAGCCTCATTGG 1767
             ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||

Db        661 TTCTGGCAATGGTTTGACGGTGTGATGGAAGTGTT-AAAAAACATCTCAAGCCTCATTGG 719

Qy       1768 AATGATGGGGCCA-TTTTGGGGTTTGTAAACAAGCAACAGG-CCCATGACCTACTGATTA 1825
             |||||||||||||| |||||||||||||||||||||||||| |||||||||||||| ||||
Db        720 AATGATGGGGCCATTTTTGGGGTTTGTAAACAAGCAACAGGCCCCATGACCTACTCATTA 779

Qy       1826 ACAAGCCAGATGGGACCTTCCT-CCTGAGATTCAGTGACTCAG-AAATTGGCGGCATCAC 1883
             |||||||||||||||||||||| |||||||||||||||||| |||||||||||||||||
Db        780 ACAAGCCAGATGGGACCTTCCTCCCTGAGATTCAGTGACTCAGAAAATTGGCGGCATCAC 839

Qy       1884 CA-TTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTTGGAAT-CTGATGCCTTTTACC 1941
              || ||||||||||||||||||||  |||||  | |||||||||| | || ||||||||||
Db        840 CATTTGCTTGGAAGTTTGATTCTTCGGAAAAGAAGTTTTGGAATCCGGAGGCCTTTTACC 899

Qy       1942 ACCAGAGAC 1950
              |||  | ||
Db        900 ACCCAAAAC 908


RESULT 10
BQ706867
LOCUS       BQ706867                 895 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8353060 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6277975
            5', mRNA sequence.
ACCESSION   BQ706867
VERSION     BQ706867.1  GI:21845766
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 895)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Mark Watson
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2465  row: d  column: 08
            High quality sequence stop: 712.
FEATURES             Location/Qualifiers
     source          1. .895
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6277975"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_113"
                     /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match            34.0%;  Score 804;  DB 5;  Length 895;
  Best Local Similarity  99.4%;  Pred. No. 2.7e-189;
  Matches  807;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1553 AAGTGCAGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGT 1612
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 AAGTGCAGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGT 60

Qy       1613 TCAACAACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGT 1672
```

```
Db         61 TCAACAACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGT 120

Qy       1673 TCAACAGGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGA 1732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAACAGGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGA 180

Qy       1733 TGGAAGTGTTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTG 1792
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGAAGTGTTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTG 240

Qy       1793 TAAACAAGCAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGA 1852
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        241 TAAACAAGCAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGA 300

Qy       1853 GATTCAGTGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAA 1912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATTCAGTGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAA 360

Qy       1913 GAATGTTTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCG 1972
             |||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db        361 GAATGTTTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCG 420

Qy       1973 ACCGCTTGGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCGCTTGGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAG 480

Qy       2033 TATACTCCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATG 2092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TATACTCCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATG 540

Qy       2093 GATACGTGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATG 2152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATACGTGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATG 600

Qy       2153 CCGGGGGCGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGG 2212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCGGGGGCGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGG 660

Qy       2213 CTCACTATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCG 2272
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTCACTATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCG 720

Qy       2273 ATCTGGAGGACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGG 2332
             |||||||||||||||||||| |||||||||||||||||||||||||||||| ||||
Db        721 ATCTGGAGGACACAATGGACGTAACGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCCATGG 780

Qy       2333 ACAGTCAGTGGATCCCGCACGCACAATCGTGA 2364
             ||||||||||||||||||||||||||||||||
Db        781 ACAGTCAGTGGATCCCGCACGCACAATCGTGA 812


RESULT 11
BQ706430
LOCUS       BQ706430                879 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8474995 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6301542
            5', mRNA sequence.
ACCESSION   BQ706430
VERSION     BQ706430.1  GI:21845329
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Mark Watson
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2517  row: j  column: 07
            High quality sequence stop: 635.
FEATURES             Location/Qualifiers
     source          1. .879
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6301542"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_113"
                     /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             33.0%;  Score 780;  DB 5;  Length 879;
  Best Local Similarity   99.0%;  Pred. No. 2.6e-183;
  Matches  805;  Conservative    0;  Mismatches    6;  Indels    2;  Gaps    2;

Qy       1553 AAGTGCAGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGT 1612
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCAGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGT 60

Qy       1613 TCAACAACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGT 1672
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAACAACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGT 120

Qy       1673 TCAACAGGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGA 1732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAACAGGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGA 180

Qy       1733 TGGAAGTGTTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTG 1792
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGAAGTGTTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTG 240

Qy       1793 TAAACAAGCAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGA 1852
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        241 TAAACAAGCAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGA 300

Qy       1853 GATTCAGTGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAA 1912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATTCAGTGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAA 360

Qy       1913 GAATGTTTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCG 1972
             |||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db        361 GAATGTTTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCG 420

Qy       1973 ACCGCTTGGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCGCTTGGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAG 480

Qy       2033 TATACTCCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATG 2092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TATACTCCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATG 540

Qy       2093 GATACGTGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATG 2152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATACGTGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATG 600

Qy       2153 CCGGGGGCGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGG 2212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCGGGGGCGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGG 660

Qy       2213 CTCACTATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCG 2272
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTCACTATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCG 720

Qy       2273 ATCTGGAGGACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGG 2332
                ||||||||||||| |||||||||| |||||||||||||||||||||||||||||| ||||||
Db        721 ATCTGGAGGACAC-ATGGACGTAACGCGGCGTGTGGAGGAGCTCCTGGGCCGGNCAATGG 779

Qy       2333 ACAGTCA-GTGGATCCCGCACGCACAATCGTGA 2364
                ||||||| |||||||||| ||||||||||||||
Db        780 ACAGTCAGGTGGATCCCGAACGCACAATCGTGA 812


RESULT 12
BX431739
LOCUS       BX431739                978 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX431739 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
            CS0DM003YD01 5-PRIME, mRNA sequence.
ACCESSION   BX431739
VERSION     BX431739.1  GI:30783021
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 978)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 8244.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAG063ZA12_CS06041_1&c=8244.r

FEATURES             Location/Qualifiers
     source          1. .978
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM003YD01"
                     /tissue_type="FETAL LIVER"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL LIVER"
                     /note="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match           32.9%;  Score 778.4;  DB 5;  Length 978;
  Best Local Similarity   89.6%;  Pred. No. 6.8e-183;
  Matches  881;  Conservative    0;  Mismatches   94;  Indels    8;  Gaps    4;

Qy        590 GTCTGAGCCGGGAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGC 649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTGAGCCGGGAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGC 60

Qy        650 GTGAGGCACAGACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCC 709
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         61 GTGAGGCACAGACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCC 120

Qy        710 TGCAGCTGCTGCGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGC 769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCAGCTGCTGCGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGC 180

Qy        770 GGCGGCAGCAGCTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGT 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAGCTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGT 240

Qy        830 CCTGGTGTGAGAAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTG 889
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGTGAGAAGTTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTG 300

Qy        890 AGCACCTCTGCCAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCA 949
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCACCTCTGCCAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCA 360

Qy        950 ACGCCACCATCACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGC 1009
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        361 ACGCCACCATCACGGACATTATCTCAGCCCTGGTGACCAGCACATTCATCATTGAGAAGC 420

Qy       1010 AGCCTCCTCAGGTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGG 1069
              |||||||||||||||||||||||||||||||||||||||||||||| || |||||||||||||
Db        421 AGCCTCCTCAGGTCCTGAAGACCCAGACCAAGTTTGCAGCCACCGTACGCCTGCTGGTGG 480

Qy       1070 GCGGGAAGCTGAACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGC 1129
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        481 GCGGGAAGCTGAACGTGCACATGAATCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGC 540

Qy       1130 AGCAGGCCAAGTCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCT 1189
              |||||||||||||||||||| || || |||||||||||||| || | |||||| ||||||
Db        541 AGCAGGCCAAGTCTCTGCTTAAAAATGAGAACACCCGCAACGAGTGCAGTGGTGAGATCC 600

Qy       1190 TGAACAACTGCTGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCA 1249
              |||||||||||||||| |||||||||||||||||| |||||||| ||||||||||||||
Db        601 TGAACAACTGCTGCGTGATGGAGTACCACCAAGCCACGGGCACCCTCAGTGCCCACTTCA 660

Qy       1250 GGAATATGTCCCTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAG 1309
              |||| ||||| |||||  | || ||| |  | ||||| ||  | |||||||| |||||||
Db        661 GGAACATGTCACTGAAGAGGATCAAGCGTGCTGACCGGCGGNGTGCAGAGTCCGTGACAG 720

Qy       1310 AAGAAAAATTTACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTT 1369
              | || || || ||| |||    |||| || ||||||||||||||  | |||||||| || |
Db        721 AGGAGAAGTTCACAGTCCCTGTTGAGTCTCAGTTCAGTGTTGGCAGCAATGAGCTTGTGT 780

Qy       1370 TTCAAGTCAAGACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACA 1429
              |||| || ||||| |||| ||| || ||||| || ||||| || || |||||||||
Db        781 TTCAGGTGAAGACTCTGTNCCTACCTGTGGTTGTCATCGTCCACGGGAGCCAGGACCCAC 840

Qy       1430 ATGCGACGGCCACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTG 1489
              |||| ||||| ||||| || |||||| |||| ||||| ||||| ||||||||||||||||
Db        841 ATGCCACGGCTACTGTGCTGTGGGAC-ATGCCTTTGCTGAGCC-GGCAGGGTGCCATTTG 898

Qy       1490 CCGTGCCTGACAAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGG 1549
              ||||||  |||||    ||  |  | || ||||  |         ||||||||  | |
Db        899 CCGTGCTGCACAAA----TGCTGGGNCCGCACTGTGTGAGCGCTCACATGAAATCAAGGC 954

Qy       1550 CCGAAGTGCAGAGCAACCGGGGC 1572
              ||  | |||| || |||||||||
Db        955 CC--AATGCAAAGAAACCGGGGC 975


RESULT 13
BQ941072
LOCUS       BQ941072                917 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8853726 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6374059
            5', mRNA sequence.
ACCESSION   BQ941072
VERSION     BQ941072.1  GI:22356550
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2552  row: g  column: 20
            High quality sequence stop: 591.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6374059"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_18"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match            32.4%;  Score 767;  DB 5;  Length 917;
  Best Local Similarity  95.5%;  Pred. No. 4.7e-180;
  Matches  843;  Conservative    0;  Mismatches   35;  Indels    5;  Gaps    5;

Qy        448 GAGCTGCGACTGGTCACGCAGGACACAGAGAATGAGTTAAAAAAGCTGCAGCAGACTCAG 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGCTGCGACTGGTCACGCAGGACACAGAGAATGAGTTAAAAAAGCTGCAGCAGACTCAG 60

Qy        508 GAGTACTTCATCATCCAGTACCAGGAGAGCCTGAGGATCCAAGCTCAGTTTGGCCCGCTG 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGTACTTCATCATCCAGTACCAGGAGAGCCTGAGGATCCAAGCTCAGTTTGGCCCGCTG 120

Qy        568 GCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGGGAGACGGCCCTCCAGCAGAAGCAGGTG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGGGAGACGGCCCTCCAGCAGAAGCAGGTG 180

Qy        628 TCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAGACACTGCAGCAGTACCGCGTGGAGCTG 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAGACACTGCAGCAGTACCGCGTGGAGCTG 240

Qy        688 CCCGAGAAGCACCAGAAGACCCTGCAGCTGCTGCGGAAGCAGCAGACCATCATCCTGGAT 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAGAAGCACCAGAAGACCCTGCAGCTGCTGCGGAAGCAGCAGACCATCATCCTGGAT 300

Qy        748 GACGAGCTGATCCAGTGGAAGCGGCGGCAGCAGCTGGCCGGGAACGGCGGGCCCCCCGAG 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACGAGCTGATCCAGTGGAAGCGGCGGCAGCAGCTGGCCGGGAACGGCGGGCCCCCCGAG 360

Qy        808 GGCAGCCTGGACGTGCTACAGTCCTGGTGTGAGAAGTTGGCCGGAGATCATCTGGCAGAAC 867
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        361 GGCAGCCTGGACGTGCTACAGTCCTGGTGTGAGAAGTTGGCCGAGATCATCTGGCAGAAC 420

Qy        868 CGGCAGCAGATCCGCAGGGCTGAGCACCTCTGCCAGCAGCTGCCCATCCCCGGCCCAGTG 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGGCAGCAGATCCGCAGGGCTGAGCACCTCTGCCAGCAGCTGCCCATCCCCGGCCCAGTG 480

Qy        928 GAGGAGATGCTGGCCGAGGTCAACGCCACCATCACGGACATTATCTCAGCCCTGGTGACC 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGGAGATGCTGGCCGAGGTCAACGCCACCATCACGGACATTATCTCAGCCCTGGTGACC 540

Qy        988 AGCACGTTCATCATTGAGAAGCAGCCTCCTCAGGTCCTGAAGACCCAGACCAAGTTTGCA 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCACGTTCATCATTGAGAAGCAGCCTCCTCAGGTCCTGAAGACCCAGACCAAGTTTGCA 600

Qy       1048 GCCACTGTGCGCCTGCTGGTGGGC-GGGAAGCTGAACGTGCACATGAACCCCCCCCAGGT 1106
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        601 GCCACTGTGCGCCTGCTGGTGGGCGGGGAAGCTGAACGTGCACATGAACCCCCCCCAGGT 660

Qy       1107 GAAGGCCACCATCATCAGTGAGCAGCAGGCCAAGTCTCTGCTCAAGAACGAGAACA-CCC 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| 
Db        661 GAAGGCCACCATCATCAGTGAGCAGCAGGCCAAGTCTCTGCTCAAGAACGAGAACACCCC 720

Qy       1166 GCAATGATTACAGTGGCGAGATCTTGAACAACTGCTGCGTCATGGAGTACCACCAAG-CC 1224
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| || 
Db        721 GCAATGATTACAGTGGCGAGATCTTGAACAACTGCTGCGTCATGGAGTACCACCAAGCCC 780

Qy       1225 ACAGGCACCCTTAGT-GCCCACTTCAGGAATATGTCCC-TGAAACGAATTAAGAGGTCAG 1282
              | |||||||||||| |||||||||||||||||||| |||||||||| | ||| | 
Db        781 CCGGGCACCCTTAGTGGCCCACTTCAGGAATATGTCCCTTGAAACGAATTTAAAGGGCCA 840

Qy       1283 ACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTTACAAT 1325
                |    |||| |    ||   | |  | ||  | |||
Db        841 GACCCGTCGGGGGGGCAAAAGTCCGGGGAACAGAAGAAAAAAT 883


RESULT 14
AL576279/c
LOCUS       AL576279                991 bp    mRNA    linear   EST 06-APR-2004
DEFINITION  AL576279 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI073YN20 3-PRIME, mRNA sequence.
ACCESSION   AL576279
VERSION     AL576279.3  GI:46249147
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 991)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 16, 2001 this sequence version replaced gi:31314566.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            8244.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI073DG10NP1&c=8244.r.
FEATURES             Location/Qualifiers
     source          1. .991
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI073YN20"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match            31.9%;  Score 753;  DB 1;  Length 991;
  Best Local Similarity  94.3%;  Pred. No. 1.5e-176;
  Matches  766;  Conservative   29;  Mismatches   15;  Indels    2;  Gaps    2;

Qy       1479 GGTGCCATTTGCCGTGCCTGACAAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACAT 1538
              ||:|||:|| |: |||| | |||:|||| |:||||||||||:||||| |||||||||||
Db        987 GGKGCCWTTGGSKGTGCTTCACAMAGTG-TKTGGCCACAGCTKTGTGA-GCGCTCAACAT 930
```



```
Qy       1539 GAAATTCAAGGCCGAAGTGCAGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCT 1598
Db        929 GAAATTCAAGCCGRAAGTSCAGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCYT 870

Qy       1599 GGCGCAGAAACTGTTCAACAACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGT 1658
Db        869 GGCGCAGAAAYTGTTCAACAACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGT 810

Qy       1659 GTCCTGGTCCCAGTTCAACAGGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATG 1718
Db        809 GTCCTGGYCCCAGTTCAACAGGGAGAATTTACCAGGMCGGAATYACACTTTCTGGCAATG 750

Qy       1719 GTTTGACGGTGTGATGGAAGTGTTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGC 1778
Db        749 GTTTGMCGGTGTGATGGMAGTGTYMAMMAAMCATCTCAAGCCTCATKGRAATGATGGGSC 690

Qy       1779 CATTTTGGGGTTTGTAAACAAGCAACAGGCCCATGACCTACTGATTAACAAGCCAGATGG 1838
Db        689 CATTTTTTKTTTTGTAAACAAGYAACAGGCCCATGACCTACTCATTAACAAGCCAGATGG 630

Qy       1839 GACCTTCCTCCTGAGATTCAGTGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTT 1898
Db        629 GACCTTCCTCCTGAGATTCAGTGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTT 570

Qy       1899 TGATTCTCAGGAAAGAATGTTTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCAT 1958
Db        569 TGATTCTCAGGAAAGAATGTTTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCAT 510

Qy       1959 CAGGTCCCTAGCCGACCGCTTGGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCG 2018
Db        509 TCGGTCCCTAGCCGACCGCTTGGGAGACTTGAATTACCTTAYCTACGTGTTTCCTGATCG 450

Qy       2019 GCCAAAAGATGAAGTATACTCCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGC 2078
Db        449 GCCAAAAGATGAAGTATACTCCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGC 390

Qy       2079 TAAAGCTGTTGATGGATACGTGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAA 2138
Db        389 TAAAGCTGTTGATGGATACGTGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAA 330

Qy       2139 CGCATCTGCAGATGCCGGGGGCGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGC 2198
Db        329 CGCATCTGCAGATGCCGGGGGCGGCAGCGCCACGTACATGGMCCASSCCCCCTCCCCAGC 270

Qy       2199 TGTGTGTCCCCAGGCTCACTATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACAC 2258
Db        269 TGTGTGTCCCCAGGCTCACTATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACAC 210

Qy       2259 CGATGGGGACTTCGATCTGGAGGACACAATGG 2290
Db        209 CGATGGGGACCTCGATCTGGAGGACACAANTG 178
```

RESULT 15
CF995633

```
LOCUS       CF995633                771 bp    mRNA    linear   EST 25-NOV-2003
DEFINITION  AGENCOURT_16110403 NIH_MGC_220 Homo sapiens cDNA clone
            IMAGE:30707033 5', mRNA sequence.
ACCESSION   CF995633
VERSION     CF995633.1  GI:38512354
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 771)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin, University of Iowa
             cDNA Library Preparation: M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM1070  row: p  column: 18
            High quality sequence stop: 661.
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30707033"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_220"
                     /note="Organ: mixed; Vector: pYX-Asc; Site_1: EcoRI;
                     Site_2: NotI; Library is oligo-dT primed and directionally
                     cloned Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according tomRNA size fraction, ligated with
                     EcoR I adaptor, digested with Not I and then cloned
                     directionally into pYX-Asc vector. Average insert size
                     0.5-1Kb. Adaptors 5'(AATTCGGCACGAGG)3' and 5'd
                     (CCTCGTGCCG)3'. 3' Linker sequence - GCGGCCGCTGAGAGCC T18.
                     Sequencing primers 3'end: T3 promoter primer 5'd
                     (ATTAACCCTCACTAAAGGGA)3'. 5' End: T7 promoter primer 5'd
                     (TAATACGACTCACTATAGGG)3'. Library was constructed in the
                     laboratory of M. Bento Soares. Average insert size 3-4kb
                     Note: this is a NIH_MGC Library."
ORIGIN
```

```
  Query Match             31.4%;  Score 743.2;  DB 7;  Length 771;
  Best Local Similarity   99.5%;  Pred. No. 3.9e-174;
  Matches  756;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       1502 AAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGA 1561
Db         13 AAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGA 72

Qy       1562 GCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAACA 1621
Db         73 GCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAACA 132

Qy       1622 GCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGG 1681
Db        133 GCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGG 192

Qy       1682 AGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTGT 1741
Db        193 AGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTGT 252

Qy       1742 TAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGC 1801
Db        253 TAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGC 312

Qy       1802 AACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGTG 1861
Db        313 AACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGTG 372

Qy       1862 ACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTT 1921
Db        373 ACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTT 432

Qy       1922 GGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTGG 1981
Db        433 GGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCGACCGCTTGG 492

Qy       1982 GAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCCA 2041
```

```
Db        493 GAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCCA 552

Qy       2042 AATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTGA 2101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 AATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTGA 612

Qy       2102 AGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGCG 2161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 AGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGCG 672

Qy       2162 GCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTATA 2221
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        673 GCAGCGCCACGTACATGGACCAGG-CCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTATA 731

Qy       2222 ACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGA 2261
              |||||||||||||||||||||||||||||||||||||||
Db        732 ACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGA 771
```

Search completed: May 17, 2005, 09:24:31
Job time : 11371 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 18:30:18 ; Search time 1828 Seconds
(without alignments)
7655.508 Million cell updates/sec

Title: US-10-052-482-162
Perfect score: 2364
Sequence: 1 atggctgtgtggatacaagc.....tcccgcacgcacaatcgtga 2364

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2364	100.0	2364	9	ADA02656	Ada02656 Human STA
2	2364	100.0	2364	10	ADB72394	Adb72394 Human STA
3	2364	100.0	2364	10	ADE95904	Ade95904 Human STA
4	2364	100.0	2782	6	ABZ35457	Abz35457 Human gen
5	2364	100.0	2782	6	ABK72299	Abk72299 Lymphona
6	2364	100.0	2782	9	ADA02655	Ada02655 Human STA
7	2364	100.0	2782	10	ADB72393	Adb72393 Human STA
8	2364	100.0	2782	10	ADE95903	Ade95903 Human DNA
9	2356	99.7	2364	12	ADP82726	Adp82726 Human STA
10	2356	99.7	2716	6	ABZ35008	Abz35008 Human gen
11	2356	99.7	2716	6	ABK84353	Abk84353 Human cDN
12	2313.2	97.9	2779	12	ADM79353	Adm79353 Human lym
13	2170	91.8	2170	6	ABS51482	Abs51482 Human cDN
14	2073.2	87.7	2776	13	ACN40207	Acn40207 Tumour-as
15	1991.2	84.2	2671	10	ADB58436	Adb58436 Toxicity-
16	1991.2	84.2	2671	10	ADB53016	Adb53016 Primary r
17	1978.4	83.7	2940	6	ABK72273	Abk72273 Lymphona
18	1976.8	83.6	2361	9	ADA02653	Ada02653 Mouse Sta
19	1976.8	83.6	2361	10	ADB72391	Adb72391 Mouse Sta
20	1976.8	83.6	2361	10	ADE95901	Ade95901 Mouse Sta
21	1976.8	83.6	2361	10	ACC79609	Acc79609 Mouse Sta
22	1976.8	83.6	2361	12	ADP82730	Adp82730 Mouse STA
23	1976.8	83.6	2574	9	ADA02652	Ada02652 Mouse Sta
24	1976.8	83.6	2574	10	ADB72390	Adb72390 Mouse Sta
25	1976.8	83.6	2574	10	ADE95900	Ade95900 Mouse DNA
26	1972	83.4	2940	12	ADM79327	Adm79327 Mouse lym
27	1894.8	80.2	2385	2	AAT33590	Aat33590 Signal tr
28	1894.8	80.2	2385	12	ADP82724	Adp82724 Human STA
29	1894.8	80.2	3120	2	AAV71029	Aav71029 Green flo
30	1894.8	80.2	3138	2	AAV71041	Aav71041 Stat5-gre
31	1894.8	80.2	3714	5	AAA89231	Aaa89231 Human sig
32	1894.8	80.2	3714	11	ADI32144	Adi32144 Human cDN
33	1894.8	80.2	4279	6	ABK84328	Abk84328 Human cDN
34	1894.8	80.2	4279	10	ADH28822	Adh28822 Human chr
35	1894.8	80.2	4298	12	ADL83236	Adl83236 Human PRO
36	1823.6	77.1	2538	12	ADQ84604	Adq84604 Human tum
37	1770	74.9	2382	10	ACC79608	Acc79608 Mouse Sta
38	1770	74.9	2382	12	ADP82728	Adp82728 Mouse STA
39	1766.8	74.7	2382	12	ADP82732	Adp82732 STAT5A 1*
40	1766.8	74.7	2382	12	ADP82735	Adp82735 STAT5A 1*
41	1741	73.6	2818	2	AAT10554	Aat10554 Sheep mam
42	506	21.4	596	12	ADN00006	Adn00006 Human STA
43	428	18.1	446	12	ADN00007	Adn00007 Human STA
44	402.2	17.0	495	5	AAA89241	Aaa89241 Human STA
45	400.2	16.9	824	6	ABQ54442	Abq54442 Human ova

ALIGNMENTS

RESULT 1
ADA02656
ID ADA02656 standard; cDNA; 2364 BP.
XX
AC ADA02656;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human STAT5B carcinoma associated coding sequence, SEQ ID NO:1174.
XX
KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003057146-A2.
XX
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-587068/55.
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
PS Claim 1; SEQ ID NO 1174; 245pp; English.
XX
CC The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism

CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed human CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2364 BP; 594 A; 636 C; 666 G; 468 T; 0 U; 0 Other;

Query Match 100.0%; Score 2364; DB 9; Length 2364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
                ||||||||||||||||||||||||
Db       2341 TGGATCCCGCACGCACAATCGTGA 2364
```

RESULT 2
ADB72394
ID ADB72394 standard; cDNA; 2364 BP.
XX
AC ADB72394;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human STAT5B cDNA.
XX
KW human; ss; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Homo sapiens.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2003-239337/23.
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 222; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a human cDNA of the invention.
XX
SQ Sequence 2364 BP; 594 A; 636 C; 666 G; 468 T; 0 U; 0 Other;

Query Match 100.0%; Score 2364; DB 10; Length 2364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60

Qy        61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120

Qy       121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180

Qy       181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240

Qy       241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300

Qy       301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360

Qy       361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420

Qy       421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480

Qy       481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540

Qy       541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600

Qy       601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660

Qy       661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720

Qy       721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780

Qy       781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840

Qy       841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
              ||||||||||||||||||||||||
Db       2341 TGGATCCCGCACGCACAATCGTGA 2364
```

RESULT 3
ADE95904
ID ADE95904 standard; DNA; 2364 BP.
XX
AC ADE95904;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human STAT5B gene coding sequence.
XX
KW cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;
KW lymphoma; breast cancer; prostate cancer; leukaemia; ds; human; STAT5B.
XX
OS Homo sapiens.
XX
PN WO2003039484-A2.
XX
PD 15-MAY-2003.
XX
PF 08-NOV-2002; 2002WO-US036071.
XX
PR 08-NOV-2001; 2001US-00052482.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2003-441462/41.
XX
PT New carcinoma associated nucleic acids and proteins, useful for screening
PT drug candidates, or for diagnosing and treating carcinomas, e.g.
PT lymphoma, breast cancer, prostate cancer or leukemia.
XX
PS Claim 1; SEQ ID NO 162; 793pp; English.
XX
CC This invention relates to novel recombinant nucleic acids for use in
CC diagnosis and treatment of cancer, especially carcinomas, as well as the
CC use of compositions in screening methods. The compositions of the
CC invention may have cytostatic activity whilst the disclosed sequences may
CC be useful for gene therapy. The carcinoma associated nucleic acids and
CC proteins are useful for diagnosing and treating carcinomas, for example
CC lymphoma, breast cancer, prostate cancer or leukaemia, or for screening
CC drug candidates or bioactive agents capable of binding to, or modulating

CC the activity of, a carcinoma associated protein. The present sequence is
CC the coding DNA sequence of the human STAT5B gene which is a carcinoma
CC associated gene of the invention.
XX
SQ Sequence 2364 BP; 594 A; 636 C; 666 G; 468 T; 0 U; 0 Other;

Query Match 100.0%; Score 2364; DB 10; Length 2364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480

Qy        481 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
```

```
Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220

Qy       2221 AACATGTACCCACACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AACATGTACCCACACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
             |||||||||||||||||||||||||
Db       2341 TGGATCCCGCACGCACAATCGTGA 2364
```

RESULT 4
ABZ35457
ID ABZ35457 standard; cDNA; 2782 BP.
XX
AC ABZ35457;
XX
DT 05-FEB-2003 (first entry)
XX
DE Human gene expression profile polynucleotide SEQ ID NO 568.
XX
KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW gene expression; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200274979-A2.
XX
PD 26-SEP-2002.
XX
PF 20-MAR-2002; 2002WO-US008456.
XX
PR 20-MAR-2001; 2001US-0276947P.
XX
PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI Wan J, Wang Y;
XX
DR WPI; 2002-740862/80.
XX
PT New gene expression profile generated from primary, endothelial,
PT epithelial, and muscle cell types, useful for identifying disease
PT pathologies involving alterations of gene expression, e.g. cancer.
XX
PS Disclosure; Page 704-705; 850pp; English.
XX
CC The invention relates to a gene expression profile comprising one or more
CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
CC is a coronary artery endothelium, umbilical artery or vein endothelium,
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC endothelium, myometrium microvascular endothelium, keratinocyte
CC epithelium, bronchial epithelium, mammary epithelium, prostate
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC small airway epithelium, renal epithelium, umbilical artery smooth
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC osteoblasts or prostate stromal cell. The gene expression profile is used
CC for determining the level of RNA expression for a sample, determining the
CC phenotype of a cell and distinguishing cell types. The gene or a protein
CC expression profile is useful in identifying disease pathologies involving
CC alterations of gene expression. The assessment of expression profiles may
CC provide meaningful information with respect to tumour type and stage,
CC treatment methods, and prognosis. The gene or protein expression profile
CC may also be used for creating microarrays. The microarray is useful for
CC genetic and physical mapping of genomes, DNA sequencing, genetic or
CC medical diagnosis, genotyping of organisms, confirming cell or tissue
CC identifications and in identifying promising antibiotics, antiviral or
CC antifungal agents
XX
SQ Sequence 2782 BP; 669 A; 739 C; 801 G; 573 T; 0 U; 0 Other;

Query Match 100.0%; Score 2364; DB 6; Length 2782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
```

```
Db        927 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
              ||||||||||||||||||||||||
Db       2487 TGGATCCCGCACGCACAATCGTGA 2510
```

```
RESULT 5
ABK72299
ID   ABK72299 standard; DNA; 2782 BP.
XX
AC   ABK72299;
XX
DT   30-JUL-2002  (first entry)
XX
DE   Lymphona associated polynucleotide #158.
XX
KW   Lymphoma associated protein; Pik3r1; GNAS; XI_(alpha_s_); NESP55; JAK1;
KW   Neurogranin; Nfr2; lymphoma; DNA vaccine; cytotoxic T-cell; animal model;
KW   leukaemia; ds.
XX
OS   Homo sapiens.
XX
PN   WO200224867-A2.
XX
PD   28-MAR-2002.
XX
PF   24-SEP-2001; 2001WO-US029798.
XX
PR   22-SEP-2000; 2000US-00668644.
PR   13-JUL-2001; 2001US-00905390.
PR   13-JUL-2001; 2001US-00905491.
XX
PA   (UYAA-) UNIV AARHUS.
XX
PI   Pedersen FS,  Sorensen AB,  Hernandez JM,  Nielsen AA,  Moving HO;
XX
DR   WPI; 2002-416279/44.
XX
PT   Novel recombinant lymphoma associated protein (LAP) such as Pik3r1, GNAS,
PT   JAK1, Neurogranin, Nfr2 proteins, useful for identifying inhibitors of
PT   LAP activity that are used for treating lymphoma.
XX
PS   Disclosure; Page 113-114; 160pp; English.
XX
```

CC The invention describes a recombinant protein (I) from a lymphoma
CC associated protein (LAP) sequence of Pik3r1, GNAS (including
CC XI_(alpha_s_), and NESP55), JAK1, Neurogranin or Nfr2 proteins. (I) is
CC useful for screening for a bioactive agent capable of binding to an LA
CC protein (LAP) which is encoded by a polynucleotide (II) and (II) is
CC useful for evaluating the effect of a candidate lymphoma drug in a
CC patient. (I) and (II) is also useful for diagnosing lymphoma involving
CC determining the expression of one or more (II), or (I) encoded by (II) in
CC a first tissue type of a first individual and comparing the expression of
CC the gene(s) from a second normal tissue type from the first individual or
CC a second unaffected individual, where a difference in the expression
CC indicates that the first individual has lymphoma. (II) is useful in the
CC diagnosis and treatment of lymphona and leukaemia. A host cell comprising
CC (II) is useful for screening drug candidates which involves providing the
CC host cell that expresses LA gene, adding a drug candidate to the cell,
CC and determining the effect of the drug candidate on the expression of the
CC LA gene. LA genes are also administered as DNA vaccines, such that
CC expression of the polypeptide encoded by the DNA vaccine, cytotoxic T-
CC cells and antibodies are induced which recognise and destroy or eliminate
CC cells expressing LA proteins. The genes are also for generating animal
CC models of lymphoma. This sequence represents a lymphona associated
CC polynucleotide described in the invention
XX
SQ Sequence 2782 BP; 669 A; 739 C; 801 G; 573 T; 0 U; 0 Other;

Query Match 100.0%; Score 2364; DB 6; Length 2782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy         1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy         1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy         1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy         1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186

Qy         2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2187 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246

Qy         2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy         2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy         2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy         2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy         2341 TGGATCCCGCACGCACAATCGTGA 2364
                ||||||||||||||||||||||||
Db         2487 TGGATCCCGCACGCACAATCGTGA 2510
```

RESULT 6
ADA02655
ID   ADA02655 standard; cDNA; 2782 BP.
XX
AC   ADA02655;
XX
DT   06-NOV-2003 (first entry)
XX
DE   Human STAT5B carcinoma associated cDNA, SEQ ID NO:1173.
XX
KW   Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW   prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003057146-A2.
XX
PD   17-JUL-2003.
XX
PF   26-DEC-2002; 2002WO-US041414.
XX
PR   26-DEC-2001; 2001US-00035832.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-587068/55.
XX
PT   New recombinant nucleic acid encoding carcinoma associated protein,
PT   useful for preparing compositions for treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 1173; 245pp; English.
XX
CC   The invention relates to recombinant carcinoma associated (CA) nucleic
CC   acid sequences from mouse and human (ADA01482-ADA03094), and to
CC   recombinant carcinoma associated proteins (CAP) encoded by them. The
CC   invention also encompasses expression vectors and host cells comprising a
CC   CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC   binds to the protein, and a biochip comprising CA nucleic acid or
CC   fragments thereof. The sequences of the invention were identified using
CC   oncogenic retroviruses, which insert into the genome of the host organism
CC   at random. Many of these do not carry transduced host oncogenes or
CC   pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC   direct consequence of the effects of proviral integration into host
CC   protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC   carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC   leukaemia) or a propensity to carcinoma by determination of the sequence
CC   of a CA gene, or by determination of CA gene expression in particular
CC   tissues. CA nucleic acids, proteins and antibodies are also useful as
CC   therapeutic agents and in screening and evaluating drug candidates. The
CC   present sequence represents a specifically claimed human CA nucleic acid
CC   sequence of the invention. Note: The complete sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2782 BP; 669 A; 739 C; 801 G; 573 T; 0 U; 0 Other;

```
  Query Match            100.0%;  Score 2364;  DB 9;  Length 2782;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2364;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy            1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy           61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy          121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy          181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy          241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy          301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy          361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy          421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy          481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          627 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy          541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826
```

```
Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
              ||||||||||||||||||||||||
Db       2487 TGGATCCCGCACGCACAATCGTGA 2510
```

```
RESULT 7
ADB72393
ID   ADB72393 standard; mRNA; 2782 BP.
XX
AC   ADB72393;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human STAT5B mRNA.
XX
KW   human; ss; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW   cancer; neoplasm; adenocarcinoma; sarcoma.
XX
OS   Homo sapiens.
XX
PN   WO2003008583-A2.
XX
PD   30-JAN-2003.
XX
PF   26-DEC-2001; 2001WO-US051291.
XX
PR   02-MAR-2001; 2001US-00798586.
PR   23-OCT-2001; 2001US-00004113.
```

PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2003-239337/23.
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 221; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a human mRNA of the invention.
XX
SQ Sequence 2782 BP; 669 A; 739 C; 801 G; 573 T; 0 U; 0 Other;

Query Match 100.0%; Score 2364; DB 10; Length 2782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TGTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
                 ||||||||||||||||||||||||
Db       2487 TGGATCCCGCACGCACAATCGTGA 2510
```

RESULT 8
ADE95903

```
ID   ADE95903 standard; DNA; 2782 BP.
XX
AC   ADE95903;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human DNA related to STAT5B gene mRNA.
XX
KW   cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;
KW   lymphoma; breast cancer; prostate cancer; leukaemia; ds; human; STAT5B.
XX
OS   Homo sapiens.
XX
PN   WO2003039484-A2.
XX
PD   15-MAY-2003.
XX
PF   08-NOV-2002; 2002WO-US036071.
XX
PR   08-NOV-2001; 2001US-00052482.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW,  Engelhard EK;
XX
DR   WPI; 2003-441462/41.
XX
PT   New carcinoma associated nucleic acids and proteins, useful for screening
PT   drug candidates, or for diagnosing and treating carcinomas, e.g.
PT   lymphoma, breast cancer, prostate cancer or leukemia.
XX
PS   Claim 1; SEQ ID NO 161; 793pp; English.
XX
CC   This invention relates to novel recombinant nucleic acids for use in
CC   diagnosis and treatment of cancer, especially carcinomas, as well as the
CC   use of compositions in screening methods. The compositions of the
CC   invention may have cytostatic activity whilst the disclosed sequences may
CC   be useful for gene therapy. The carcinoma associated nucleic acids and
CC   proteins are useful for diagnosing and treating carcinomas, for example
CC   lymphoma, breast cancer, prostate cancer or leukaemia, or for screening
CC   drug candidates or bioactive agents capable of binding to, or modulating
CC   the activity of, a carcinoma associated protein. The present sequence is
CC   a DNA sequence which represents the mRNA derived from the human STAT5B
CC   gene which is a carcinoma associated gene of the invention.
XX
SQ   Sequence 2782 BP; 669 A; 739 C; 801 G; 573 T; 0 U; 0 Other;
```

Query Match 100.0%; Score 2364; DB 10; Length 2782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 266

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 326

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 386

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 446

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 506

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 566

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 626

Qy        481 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 686

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 746

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 806

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 866

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
```

```
Db        867 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 926

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840

Db        927 CTGGCCGGGAACGGCGGGCCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 986

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900

Db        987 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1046

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960

Db       1047 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1106

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020

Db       1107 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1166

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080

Db       1167 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1226

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140

Db       1227 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1286

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200

Db       1287 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1346

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260

Db       1347 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1406

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320

Db       1407 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1466

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380

Db       1467 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1526

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440

Db       1527 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1586

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500

Db       1587 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1646

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560

Db       1647 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1706

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620

Db       1707 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1766

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680

Db       1767 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1826

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740

Db       1827 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1886

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800

Db       1887 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1946

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860

Db       1947 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2006

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920

Db       2007 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2066

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980

Db       2067 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 2126

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040

Db       2127 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2186

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100

Db       2187 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2246

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160

Db       2247 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2306

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220

Db       2307 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2366

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280

Db       2367 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2426

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340

Db       2427 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2486

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364

Db       2487 TGGATCCCGCACGCACAATCGTGA 2510
```

```
RESULT 9
ADP82726
ID   ADP82726 standard; DNA; 2364 BP.
XX
AC   ADP82726;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Human STAT5B protein encoding sequence.
XX
KW   pluripotency; signal transducer and activator of transcription 5; STAT5;
KW   Cytostatic; Dermatological; Antiinflammatory; Antithyroid; Antidiabetic;
KW   Hepatotropic; Nephrotropic; Cardiant; prophylaxis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2364
FT                   /*tag= a
FT                   /product= "STAT5B"
XX
PN   WO2004042040-A1.
XX
PD   21-MAY-2004.
XX
PF   23-OCT-2003; 2003WO-JP013579.
XX
PR   08-NOV-2002; 2002JP-00326190.
XX
PA   (REPR-) REPROCELL INC.
XX
PI   Nakauchi H,  Iwama A;
XX
DR   WPI; 2004-411515/38.
```

```
DR   P-PSDB; ADP82727.
XX
PT   New composition comprising active STAT5 or an agent capable of activating
PT   STAT5, useful for maintaining the expansion, pluripotency, or self-
PT   replication ability of a stem cell.
XX
PS   Claim 5; SEQ ID NO 3; 249pp; English.
XX
CC   The present invention relates to a composition for maintaining the
CC   expansion, pluripotency, or self-replication ability of a stem cell
CC   comprises active signal transducer and activator of transcription 5
CC   (STAT5), an agent capable of activating STAT5, and a nucleic acid
CC   molecule encoding active STAT5. The STAT5, active STAT5, or an agent
CC   capable of activating STAT5 is useful for maintaining the expansion,
CC   pluripotency, or self-replication ability of a stem cell. The cell
CC   obtained by treating a stem cell with active STAT5, or STAT5 and an agent
CC   capable of activating the STAT5 is useful for the treatment or
CC   prophylaxis of a disease or disorder in need of a stem cell or a
CC   differentiated cell derived from it, or for preparing a medicament for
CC   the treatment or prophylaxis of the disease or disorder. Diseases include
CC   atrichia, melanoma, cutis matignant lymphoma, hemangiosarcoma,
CC   histiocytosis, hydroa, pustulosis, dermatitis, or eczema. It also
CC   includes hypothalamus/hypophysis diseases, thyroid gland diseases,
CC   accessory thyroid gland diseases, lipid metabolism abnormality, inborn
CC   error of metabolism, hyperbilirubinemia, hyperbilirubinuria, mast cell
CC   deficiency, diabetes insipidus, dwarfism, or Wolman's disease. It further
CC   includes liver cirrhosis, pancreatitis, pancreas cancer, kidney diseases,
CC   bladder diseases, male/female genital organ disease, heart failure, or
CC   myocardial infarction. The present sequence represents STAT5B protein
CC   encoding sequence.
XX
SQ   Sequence 2364 BP; 593 A; 637 C; 665 G; 469 T; 0 U; 0 Other;

  Query Match             99.7%;  Score 2356;  DB 12;  Length 2364;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2359;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        661 ACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGTTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
                |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1801 CAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
                |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCGACCGCTTG 1980

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
                |||||||||||||||||||||||||
Db       2341 TGGATCCCGCACGCACAATCGTGA 2364
```

```
RESULT 10
ABZ35008
ID   ABZ35008 standard; cDNA; 2716 BP.
XX
AC   ABZ35008;
XX
DT   05-FEB-2003  (first entry)
XX
DE   Human gene expression profile polynucleotide SEQ ID NO 120.
XX
KW   Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW   bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW   tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW   gene expression; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200274979-A2.
XX
PD   26-SEP-2002.
XX
PF   20-MAR-2002; 2002WO-US008456.
XX
PR   20-MAR-2001; 2001US-0276947P.
XX
PA   (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI   Wan J,  Wang Y;
XX
DR   WPI; 2002-740862/80.
XX
PT   New gene expression profile generated from primary, endothelial,
PT   epithelial, and muscle cell types, useful for identifying disease
PT   pathologies involving alterations of gene expression, e.g. cancer.
XX
PS   Claim 3; Page 305-306; 850pp; English.
XX
CC   The invention relates to a gene expression profile comprising one or more
CC   genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
CC   is a coronary artery endothelium, umbilical artery or vein endothelium,
CC   aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC   endothelium, myometrium microvascular endothelium, keratinocyte
CC   epithelium, bronchial epithelium, mammary epithelium, prostate
CC   epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC   small airway epithelium, renal epithelium, umbilical artery smooth
CC   muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC   dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC   aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC   bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC   osteoblasts or prostate stromal cell. The gene expression profile is used
CC   for determining the level of RNA expression for a sample, determining the
CC   phenotype of a cell and distinguishing cell types. The gene or a protein
CC   expression profile is useful in identifying disease pathologies involving
CC   alterations of gene expression. The assessment of expression profiles may
CC   provide meaningful information with respect to tumour type and stage,
CC   treatment methods, and prognosis. The gene or protein expression profile
CC   may also be used for creating microarrays. The microarray is useful for
CC   genetic and physical mapping of genomes, DNA sequencing, genetic or
CC   medical diagnosis, genotyping of organisms, confirming cell or tissue
CC   identifications and in identifying promising antibiotics, antiviral or
CC   antifungal agents
XX
SQ   Sequence 2716 BP; 662 A; 707 C; 760 G; 587 T; 0 U; 0 Other;

  Query Match             99.7%;  Score 2356;  DB 6;  Length 2716;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2359;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        12 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAGGCG 71

Qy        61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        72 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 131

Qy       121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       132 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 191

Qy       181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       192 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 251

Qy       241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       252 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 311

Qy       301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       312 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 371

Qy       361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        372 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 431
Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
Db        432 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 491
Qy        481 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
Db        492 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 551
Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
Db        552 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 611
Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
Db        612 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 671
Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
Db        672 ACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 731
Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
Db        732 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 791
Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
Db        792 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 851
Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
Db        852 AAGTTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 911
Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
Db        912 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 971
Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
Db        972 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1031
Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
Db       1032 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1091
Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
Db       1092 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1151
Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
Db       1152 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1211
Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
Db       1212 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1271
Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
Db       1272 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1331
Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
Db       1332 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1391
Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
Db       1392 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1451
Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
Db       1452 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1511

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
Db       1512 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1571
Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
Db       1572 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1631
Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
Db       1632 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1691
Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
Db       1692 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1751
Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
Db       1752 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1811
Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
Db       1812 CAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1871
Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
Db       1872 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1931
Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
Db       1932 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCCGGTCCCTAGCCGACCGCTTG 1991
Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
Db       1992 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2051
Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
Db       2052 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2111
Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
Db       2112 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2171
Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
Db       2172 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2231
Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
Db       2232 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2291
Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
Db       2292 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2351
Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
Db       2352 TGGATCCCGCACGCACAATCGTGA 2375
```

```
RESULT 11
ABK84353
ID   ABK84353 standard; cDNA; 2716 BP.
XX
AC   ABK84353;
XX
DT   14-AUG-2002  (first entry)
XX
DE   Human cDNA differentially expressed in granulocytic cells #924.
XX
KW   Human; ss; granulocytic cell; DNA chip; bacterial infection;
```

KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US030821.
XX
PR 03-OCT-2000; 2000US-0237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX
PS Claim 1; SEQ ID NO 924; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2716 BP; 662 A; 707 C; 760 G; 587 T; 0 U; 0 Other;

Query Match 99.7%; Score 2356; DB 6; Length 2716;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2359; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db         12 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAGGCG 71

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 131

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 191

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 251

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 311

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 371

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 431

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 491

Qy        481 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 551

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 611

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 671

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
                |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        672 ACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 731

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 791

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 851

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
                |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 AAGTTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 911

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 971

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1031

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1032 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1091

Qy       1081 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1092 AACGTGCACATGAACCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1151

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1152 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1211

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1212 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1271

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1272 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1331

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1332 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1391

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1392 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1451

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1452 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1511

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1512 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1571

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1572 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1631

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1632 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1691

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1692 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1751

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1752 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1811

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1812 CAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1871

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1872 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1931

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1932 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCCGGTCCCTAGCCGACCGCTTG 1991

Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1992 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2051

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2052 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2111

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2112 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2171

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2172 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2231

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2232 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2291

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2292 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2351

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
              |||||||||||||||||||||||||
Db       2352 TGGATCCCGCACGCACAATCGTGA 2375
```

RESULT 12
ADM79353

```
ID   ADM79353 standard; cDNA; 2779 BP.
XX
AC   ADM79353;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Human lymphoma associated, LA, contig #3.
XX
KW   ss; lymphoma associated protein; LAP; GNAS1; Pik3r1; G protein XI_as;
KW   NESP55; HIPK1; JAKI; neurogranin; Nrf2; lymphoma; leukaemia;
KW   gene therapy; DNA vaccine; wound; inflammation; human;
KW   expressed sequence tag; EST.
XX
OS   Homo sapiens.
XX
PN   US2003224460-A1.
XX
PD   04-DEC-2003.
XX
PF   24-SEP-2001; 2001US-00963131.
XX
PR   22-SEP-2000; 2000US-00668644.
PR   13-JUL-2001; 2001US-00905390.
PR   13-JUL-2001; 2001US-00905491.
XX
PA   (PEDE/) PEDERSEN F S.
PA   (SORE/) SORENSEN A B.
PA   (HERN/) HERNANDEZ J M.
PA   (NIEL/) NIELSEN A A.
PA   (MOVI/) MOVING H.
XX
PI   Pedersen FS,  Sorensen AB,  Hernandez JM,  Nielsen AA,  Moving H;
XX
DR   WPI; 2004-033970/03.
XX
PT   Novel recombinant protein comprising lymphoma associated protein, useful
PT   in treating lymphoma and leukemia.
XX
PS   Disclosure; SEQ ID NO 158; 99pp; English.
XX
CC   The invention relates to a recombinant protein comprising lymphoma
CC   associated protein (LAP), chosen from a fully defined amino acid sequence
CC   of human GNAS1, mouse Pik3r1, human Pik3r1, mouse G protein XI_as, human
CC   G protein XI_as, mouse NESP55, human NESP55, mouse GNAS1, mouse HIPK1,
CC   human HIPK1, mouse JAKI, human JAKI, mouse neurogranin, human
CC   neurogranin, mouse Nrf2 and human Nrf2. LAP is useful for screening a
CC   bioactive agent capable of binding to LAP. LAP is useful for screening
CC   bioactive agents capable for modulating the activity of LAP. Lymphoma
CC   associated (LA) gene is useful for diagnosing lymphoma. LAP is useful for
CC   treating lymphoma. LA gene is useful for diagnosing lymphomas or a
CC   propensity to lymphomas. LA gene is useful for determining LA gene copy
CC   number. LA genes are used as probe to determine chromosome location of
CC   the LA gene. LA gene and LAP are useful in treating lymphoma and
CC   leukaemia. LA gene is useful in gene therapy, useful as DNA vaccine. LA
CC   gene is useful in generating animal models of lymphoma, which is useful
CC   in screening bioactive molecules to treat lymphoma. LAP is useful in
CC   treating wounds and inflammation. LAP is useful to generate polyclonal
```

CC and monoclonal antibodies. The present sequence represents a human
CC lymphoma associated, LA, contig.
XX
SQ Sequence 2779 BP; 675 A; 735 C; 801 G; 559 T; 0 U; 9 Other;

Query Match 97.9%; Score 2313.2; DB 12; Length 2779;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2349; Conservative 1; Mismatches 12; Indels 4; Gaps 3;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        147 ATGGCTGTGTGGATACAAGCTCAGNAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 206

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTG--AAA 118
              ||||||||||||||||||||||||||||||||||| |||||||||||||||  |||
Db        207 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTA-TTATCCCAGTGGATTGAAAAA 265

Qy        119 GCCAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGC 178
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        266 GCCAAGCATGGGACTCAGTAGATC-TGATAATCCACAGGAGAACATTAAGGCCACCCAGC 324

Qy        179 TCCTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATG 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 TCCTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATG 384

Qy        239 GGTTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACC 298
              |||||| |||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db        385 GGTTTTNACTGNAGATGAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACC 444

Qy        299 GCTGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 GCTGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGG 504

Qy        359 TCCGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGA 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 TCCGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGA 564

Qy        419 AACACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGA 478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 AACACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGA 624

Qy        479 ATGAGTTAAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCC 538
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        625 ATGAGTTAAAAAAAGCTGCAGNAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCC 684

Qy        539 TGAGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCC 598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 TGAGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCC 744

Qy        599 GGGAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCAC 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 GGGAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCAC 804

Qy        659 AGACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGC 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 AGACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGC 864

Qy        719 TGCGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGC 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 TGCGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGC 924

Qy        779 AGCTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTG 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 AGCTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTG 984

Qy        839 AGAAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCT 898
              |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 AGAAGAAAGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCT 1044

Qy        899 GCCAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCA 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GCCAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCA 1104

Qy        959 TCACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTC 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1105 TCACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTC 1164

Qy        1019 AGGTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGC 1078
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1165 AGGTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGC 1224

Qy        1079 TGAACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCA 1138
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1225 TGAACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCA 1284

Qy        1139 AGTCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACT 1198
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1285 AGTCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACT 1344

Qy        1199 GCTGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGT 1258
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1345 GCTGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGT 1404

Qy        1259 CCCTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAAT 1318
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1405 CCCTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAAT 1464

Qy        1319 TTACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCA 1378
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1465 TTACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCA 1524

Qy        1379 AGACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGG 1438
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1525 AGACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGG 1584

Qy        1439 CCACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTG 1498
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1585 CCACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTG 1644

Qy        1499 ACAAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGC 1558
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1645 ACAAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGC 1704

Qy        1559 AGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACA 1618
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1705 AGAGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACA 1764

Qy        1619 ACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACA 1678
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1765 ACAGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACA 1824

Qy        1679 GGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAG 1738
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1825 GGGAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAG 1884

Qy        1739 TGTTAAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACA 1798
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1885 TGTTAAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACA 1944

Qy        1799 AGCAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCA 1858
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1945 AGCAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCA 2004

Qy        1859 GTGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGT 1918
               |||||||||||||||||||||||||||||||||||||||: ||||||||||||||||||||
Db        2005 GTGACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGWNTGATTCTCAGGAAAGAATGT 2064

Qy        1919 TTTGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCT 1978
               |||||||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Db        2065 TTTGGAATCTGATGCCNTNTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCT 2124

Qy        1979 TGGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACT 2038
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2125 TGGGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACT 2184

Qy        2039 CCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACG 2098
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2185 CCAAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACG 2244

Qy       2099 TGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGG 2158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2245 TGAAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGG 2304

Qy       2159 GCGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACT 2218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2305 GCGGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACT 2364

Qy       2219 ATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGG 2278
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       2365 ATAACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACNTCGATCTGG 2424

Qy       2279 AGGACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTC 2338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2425 AGGACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTC 2484

Qy       2339 AGTGGATCCCGCACGCACAATCGTGA 2364
              |||||||||||||||||||||||||
Db       2485 AGTGGATCCCGCACGCACAATCGTGA 2510


RESULT 13
ABS51482
ID   ABS51482 standard; cDNA; 2170 BP.
XX
AC   ABS51482;
XX
DT   21-OCT-2002  (first entry)
XX
DE   Human cDNA encoding prey protein for Shigella ospC1 #7.
XX
KW   ss; gene; prey protein; ospB; ospD1; ipaD; ipaC; ipaH9.8; ospG; ospC1;
KW   Shigella; shigellosis; bacillary dysentery; antibacterial;
KW   yeast two-hybrid system; protein-protein interaction; SID;
KW   selected interacting domain; human.
XX
OS   Homo sapiens.
XX
PN   WO200257303-A2.
XX
PD   25-JUL-2002.
XX
PF   11-JAN-2002; 2002WO-EP000777.
XX
PR   12-JAN-2001; 2001US-0261130P.
XX
PA   (HYBR-) HYBRIGENICS.
XX
PI   Legrain P;
XX
DR   WPI; 2002-599706/64.
DR   P-PSDB; ABG70089.
XX
PT   New complex of protein-protein interactions between a bait Shigella
PT   flexneri polypeptide and a prey mammalian or human placenta polypeptide
PT   for treating or preventing bacillary dysentery in a mammal or human.
XX
PS   Claim 8; Page 74-75; 162pp; English.
XX
CC   The invention relates to a complex of protein-protein interactions
CC   between a Shigella flexneri polypeptide (e.g. ospB, ospD1, ipaD, ipaC,
CC   ipaH9.8, ospG and ospC1) and a mammalian polypeptide defined in the
CC   specification. The complexes are formed using the yeast two-hybrid
CC   system. Also included are (1) a recombinant host cell expressing the
CC   interactions between the Shigella flexneri polypeptide and a mammalian
CC   polypeptide defined in the specification; (2) selecting a modulating
CC   compound that inhibits or activates the protein-protein interactions; (3)
CC   a modulating compound obtained from the method of (2); (4) a SID
CC   (selected interacting domain) polypeptide or its fragment or variant
CC   comprising the human polypeptides appearing as ABG70042-ABG70242; (5) a
CC   SID polynucleotide or its fragment or variant comprising encoding the
CC   above polypeptides a vector comprising (5); (6) a recombinant host cell
CC   containing the vector; and (10) a protein chip comprising Shigella
CC   flexneri polypeptide and a mammalian polypeptide defined in the
CC   specification. A pharmaceutical composition comprising the compound,
CC   polypeptide or polynucleotide is useful for treating or preventing
CC   shigellosis (bacillary dysentery) in a human or mammal. The present
CC   sequence encodes a human prey protein isolated by the yeast two-hybrid
CC   assay, forming a complex of the invention with a shigella protein
XX
SQ   Sequence 2170 BP; 542 A; 590 C; 615 G; 423 T; 0 U; 0 Other;

  Query Match             91.8%;  Score 2170;  DB 6;  Length 2170;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2170;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        195 GCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGGTTTTTACTGAAGAT 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGGTTTTTACTGAAGAT 60

Qy        255 CAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGCTGCCCCATGGAGCT 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGCTGCCCCATGGAGCT 120

Qy        315 GGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTCCGAGAAGCCAACAA 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTCCGAGAAGCCAACAA 180

Qy        375 TGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAACACCTCCAGATCAA 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAACACCTCCAGATCAA 240

Qy        435 CCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAATGAGTTAAAAAAGCT 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAATGAGTTAAAAAAGCT 300

Qy        495 GCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTGAGGATCCAAGCTCA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTGAGGATCCAAGCTCA 360

Qy        555 GTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGGGAGACGGCCCTCCA 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGGGAGACGGCCCTCCA 420

Qy        615 GCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAGACACTGCAGCAGTA 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAGACACTGCAGCAGTA 480

Qy        675 CCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTGCGGAAGCAGCAGAC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTGCGGAAGCAGCAGAC 540

Qy        735 CATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAGCTGGCCGGGAACGG 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAGCTGGCCGGGAACGG 600

Qy        795 CGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAGAAGTTGGCGGAGAT 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAGAAGTTGGCGGAGAT 660

Qy        855 CATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGCCAGCAGCTGCCCAT 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGCCAGCAGCTGCCCAT 720

Qy        915 CCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATCACGGACATTATCTC 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATCACGGACATTATCTC 780

Qy        975 AGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAGGTCCTGAAGACCCA 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAGGTCCTGAAGACCCA 840

Qy       1035 GACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTGAACGTGCACATGAA 1094
```

```
Db        841 GACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTGAACGTGCACATGAA 900

Qy       1095 CCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAGTCTCTGCTCAAGAA 1154
Db        901 CCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAGTCTCTGCTCAAGAA 960

Qy       1155 CGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGCTGCGTCATGGAGTA 1214
Db        961 CGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGCTGCGTCATGGAGTA 1020

Qy       1215 CCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCCCTGAAACGAATTAA 1274
Db       1021 CCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCCCTGAAACGAATTAA 1080

Qy       1275 GAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTTACAATCCTGTTTGA 1334
Db       1081 GAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTTACAATCCTGTTTGA 1140

Qy       1335 ATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAGACCCTGTCCCTGCC 1394
Db       1141 ATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAGACCCTGTCCCTGCC 1200

Qy       1395 AGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCCACTGTTCTCTGGGA 1454
Db       1201 AGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCCACTGTTCTCTGGGA 1260

Qy       1455 CAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGCTGTGGCC 1514
Db       1261 CAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGACAAAGTGCTGTGGCC 1320

Qy       1515 ACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACCGGGGCCT 1574
Db       1321 ACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAGAGCAACCGGGGCCT 1380

Qy       1575 GACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCAGCCACCT 1634
Db       1381 GACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAACAGCAGCAGCCACCT 1440

Qy       1635 GGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGGAGAATTTACCAGG 1694
Db       1441 GGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGGGAGAATTTACCAGG 1500

Qy       1695 ACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTGTTAAAAAAACATCT 1754
Db       1501 ACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTGTTAAAAAAACATCT 1560

Qy       1755 CAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGCAACAGGCCCATGA 1814
Db       1561 CAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAGCAACAGGCCCATGA 1620

Qy       1815 CCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGTGACTCAGAAATTGG 1874
Db       1621 CCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGTGACTCAGAAATTGG 1680

Qy       1875 CGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTTGGAATCTGATGCC 1934
Db       1681 CGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTTTGGAATCTGATGCC 1740

Qy       1935 TTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTGGGAGACTTGAATTA 1994
Db       1741 TTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTGGGAGACTTGAATTA 1800

Qy       1995 CCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCCAAATACTACACACC 2054
Db       1801 CCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCCAAATACTACACACC 1860

Qy       2055 AGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTGAAGCCACAGATCAA 2114
Db       1861 AGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTGAAGCCACAGATCAA 1920

Qy       2115 GCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGCGGCAGCGCCACGTA 2174
```

```
Db       1921 GCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGCGGCAGCGCCACGTA 1980

Qy       2175 CATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTATAACATGTACCCACA 2234
Db       1981 CATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTATAACATGTACCCACA 2040

Qy       2235 GAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAGGACACAATGGACGT 2294
Db       2041 GAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAGGACACAATGGACGT 2100

Qy       2295 AGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAGTGGATCCCGCACGC 2354
Db       2101 AGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAGTGGATCCCGCACGC 2160

Qy       2355 ACAATCGTGA 2364
Db       2161 ACAATCGTGA 2170
```

RESULT 14
ACN40207
ID ACN40207 standard; cDNA; 2776 BP.
XX
AC ACN40207;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA326348, SEQ ID NO:4849.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 4849; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with

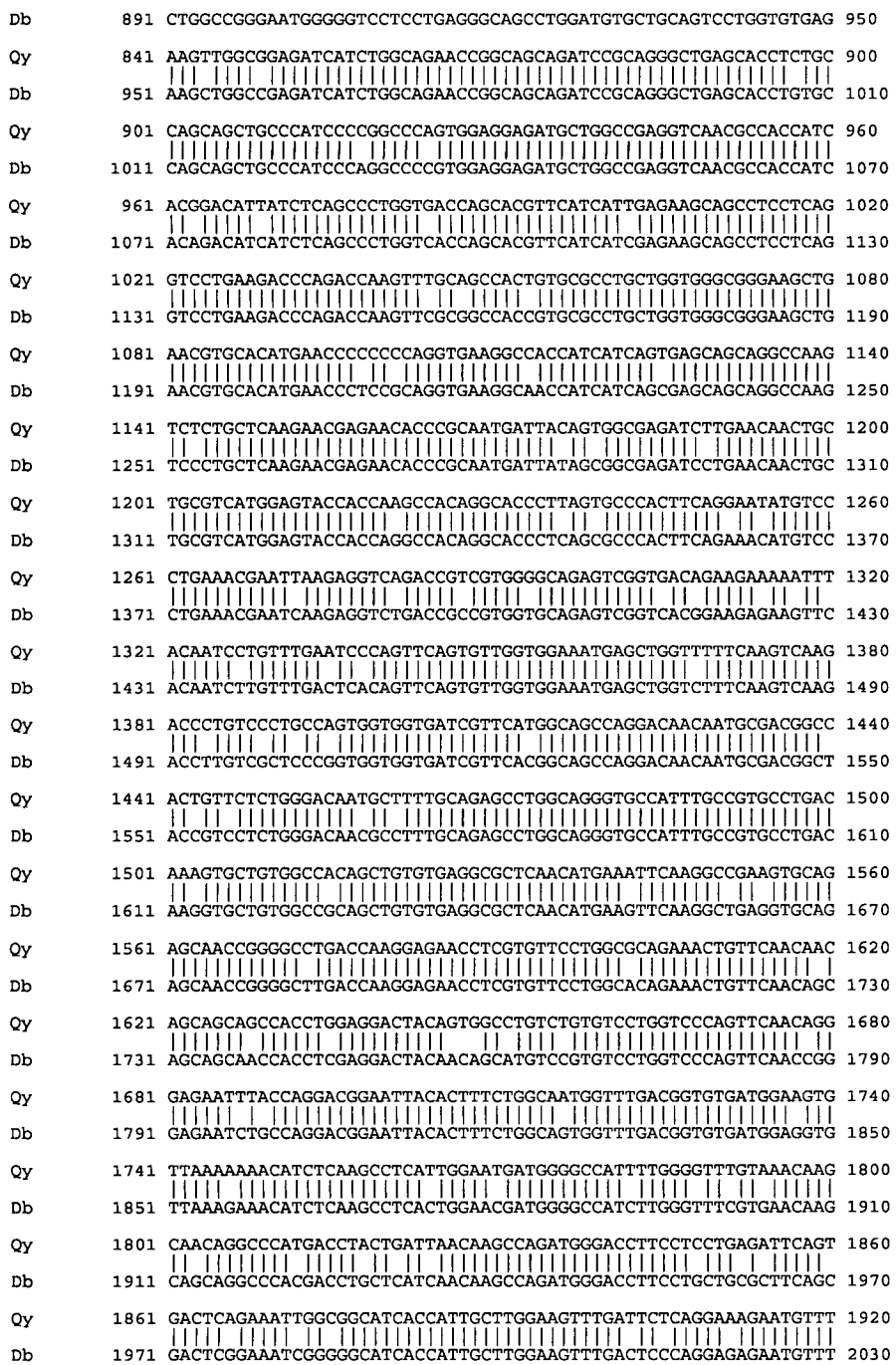
```
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT nucleic acid of the invention
XX
SQ   Sequence 2776 BP; 607 A; 684 C; 734 G; 513 T; 0 U; 238 Other;

  Query Match            87.7%;  Score 2073.2;  DB 13;  Length 2776;
  Best Local Similarity  88.9%;  Pred. No. 0;
  Matches 2101;  Conservative    0;  Mismatches  256;  Indels    7;  Gaps    1;

Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 207

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 267

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 327

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 387

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 447

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 507

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 567

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 627

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 687

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 747

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 807

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        808 ACACTGCAGCAGTACCGCGTGGAGCTGGCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 867

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 927

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 987

Qy        841 AAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 900
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 AAGTTGGCCGAGATCATCTGGCAGAACCGGCAGCAGATCCGCAGGGCTGAGCACCTCTGC 1047

Qy        901 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 CAGCAGCTGCCCATCCCCGGCCCAGTGGAGGAGATGCTGGCCGAGGTCAACGCCACCATC 1107

Qy        961 ACGGACATTATCTCAGCCCTGGTGACCAGCACGTTCATCATTGAGAAGCAGCCTCCTCAG 1020
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1108 ACGGACATTATCTCAGCCCTGGTGACCAGCACATTCATCATTGAGAAGCAGCCTCCTCAG 1167

Qy       1021 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACTGTGCGCCTGCTGGTGGGCGGGAAGCTG 1080
              |||||||||||||||||||||||||||||| || ||||||||||||||||||||||||||
Db       1168 GTCCTGAAGACCCAGACCAAGTTTGCAGCCACCGTACGCCTGCTGGTGGGCGGGAAGCTG 1227

Qy       1081 AACGTGCACATGAACCCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1140
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1228 AACGTGCACATGAATCCCCCCCAGGTGAAGGCCACCATCATCAGTGAGCAGCAGGCCAAG 1287

Qy       1141 TCTCTGCTCAAGAACGAGAACACCCGCAATGATTACAGTGGCGAGATCTTGAACAACTGC 1200
              |||||||| || || |||||||||||||          ||| |||||| ||||||||||
Db       1288 TCTCTGCTTAAAAATGAGAACACCCGCAANNNNNNNNNNTGGTGAGATCCTGAACAACTGC 1347

Qy       1201 TGCGTCATGGAGTACCACCAAGCCACAGGCACCCTTAGTGCCCACTTCAGGAATATGTCC 1260
              ||||| |||||||||||||||||||| ||||||||       ||
Db       1348 TGCGTGATGGAGTACCACCAAGCCACGGGCACCCT-------CANNNNNNNNNNNNNNN 1400

Qy       1261 CTGAAACGAATTAAGAGGTCAGACCGTCGTGGGGCAGAGTCGGTGACAGAAGAAAAATTT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1401 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1460

Qy       1321 ACAATCCTGTTTGAATCCCAGTTCAGTGTTGGTGGAAATGAGCTGGTTTTTCAAGTCAAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1461 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1520

Qy       1381 ACCCTGTCCCTGCCAGTGGTGGTGATCGTTCATGGCAGCCAGGACAACAATGCGACGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1521 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1580

Qy       1441 ACTGTTCTCTGGGACAATGCTTTTGCAGAGCCTGGCAGGGTGCCATTTGCCGTGCCTGAC 1500
                                                 ||||||||||||||||||||||||||
Db       1581 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGCAGGGTGCCATTTGCCGTGCCTGAC 1640

Qy       1501 AAAGTGCTGTGGCCACAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1560
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1641 AAAGTGCTGTGGCCGCAGCTGTGTGAGGCGCTCAACATGAAATTCAAGGCCGAAGTGCAG 1700

Qy       1561 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1701 AGCAACCGGGGCCTGACCAAGGAGAACCTCGTGTTCCTGGCGCAGAAACTGTTCAACAAC 1760

Qy       1621 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCTGTGTCCTGGTCCCAGTTCAACAGG 1680
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1761 AGCAGCAGCCACCTGGAGGACTACAGTGGCCTGTCCGTGTCCTGGTCCCAGTTCAACAGG 1820

Qy       1681 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1821 GAGAATTTACCAGGACGGAATTACACTTTCTGGCAATGGTTTGACGGTGTGATGGAAGTG 1880

Qy       1741 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1881 TTAAAAAAACATCTCAAGCCTCATTGGAATGATGGGGCCATTTTGGGGTTTGTAAACAAG 1940

Qy       1801 CAACAGGCCCATGACCTACTGATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 1860
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1941 CAACAGGCCCATGACCTACTCATTAACAAGCCAGATGGGACCTTCCTCCTGAGATTCAGT 2000

Qy       1861 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 GACTCAGAAATTGGCGGCATCACCATTGCTTGGAAGTTTGATTCTCAGGAAAGAATGTTT 2060

Qy       1921 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATCAGGTCCCTAGCCGACCGCTTG 1980
              |||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db       2061 TGGAATCTGATGCCTTTTACCACCAGAGACTTCTCCATTCGGTCCCTAGCCGACCGCTTG 2120
```

```
Qy       1981 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GGAGACTTGAATTACCTTATCTACGTGTTTCCTGATCGGCCAAAAGATGAAGTATACTCC 2180

Qy       2041 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 AAATACTACACACCAGTTCCCTGCGAGTCTGCTACTGCTAAAGCTGTTGATGGATACGTG 2240

Qy       2101 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 AAGCCACAGATCAAGCAAGTGGTCCCTGAGTTTGTGAACGCATCTGCAGATGCCGGGGGC 2300

Qy       2161 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2301 GGCAGCGCCACGTACATGGACCAGGCCCCCTCCCCAGCTGTGTGTCCCCAGGCTCACTAT 2360

Qy       2221 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2361 AACATGTACCCACAGAACCCTGACTCAGTCCTTGACACCGATGGGGACTTCGATCTGGAG 2420

Qy       2281 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2421 GACACAATGGACGTAGCGCGGCGTGTGGAGGAGCTCCTGGGCCGGCCAATGGACAGTCAG 2480

Qy       2341 TGGATCCCGCACGCACAATCGTGA 2364
             |||||||||||||||||||||||||
Db       2481 TGGATCCCGCACGCACAATCGTGA 2504
```

RESULT 15
ADB58436
ID ADB58436 standard; DNA; 2671 BP.
XX
AC ADB58436;
XX
DT 04-DEC-2003 (first entry)
XX
DE Toxicity-related gene, SEQ ID 3462.
XX
KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW drug screening; toxicity assay; ds.
XX
OS Unidentified.
XX
PN WO2003064624-A2.
XX
PD 07-AUG-2003.
XX
PF 31-JAN-2003; 2003WO-US003194.
XX
PR 31-JAN-2002; 2002US-00060087.
PR 15-MAR-2002; 2002US-0364045P.
PR 15-MAR-2002; 2002US-0364055P.
PR 30-DEC-2002; 2002US-0436643P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
DR WPI; 2003-689530/65.
XX
PT Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to
PT the compound.
XX
PS Claim 1; SEQ ID NO 3462; 1156pp; English.
XX
CC The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2671 BP; 638 A; 739 C; 776 G; 518 T; 0 U; 0 Other;

Query Match 84.2%; Score 1991.2; DB 10; Length 2671;
Best Local Similarity 90.4%; Pred. No. 0;
Matches 2138; Conservative 0; Mismatches 223; Indels 3; Gaps 1;

```
Qy          1 ATGGCTGTGTGGATACAAGCTCAGCAGCTCCAAGGAGAAGCCCTTCATCAGATGCAAGCG 60
              |||||  |||||||||| |||||||||||||| || || |||||||||||||||| ||
Db        111 ATGGCAATGTGGATACAGGCTCAGCAGCTCCAGGGCGATGCCCTTCATCAGATGCAGGCC 170

Qy         61 TTATATGGCCAGCATTTTCCCATTGAGGTGCGGCATTATTTATCCCAGTGGATTGAAAGC 120
              || || |||||||||| ||||| |||||||| |||||||||||||||||||| |||||
Db        171 TTGTACGGCCAGCATTTCCCCATCGAGGTGCGACATTATTTATCCCAGTGGATCGAAAGT 230

Qy        121 CAAGCATGGGACTCAGTAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 180
              ||||| ||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        231 CAAGCCTGGGACTCAATAGATCTTGATAATCCACAGGAGAACATTAAGGCCACCCAGCTC 290

Qy        181 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAGCACCAGGTGGGGGAAGATGGG 240
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        291 CTGGAGGGCCTGGTGCAGGAGCTGCAGAAGAAGGCAGAACACCAGGTGGGGGAAGATGGG 350

Qy        241 TTTTTACTGAAGATCAAGCTGGGGCACTATGCCACACAGCTCCAGAACACGTATGACCGC 300
              ||||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        351 TTTTTGCTGAAGATCAAGCTGGGGCACTATGCCACGCAGCTCCAGAACACGTATGACCGC 410

Qy        301 TGCCCCATGGAGCTGGTCCGCTGCATCCGCCATATATTGTACAATGAACAGAGGTTGGTC 360
              |||||||||||||||||| ||||| |||| || ||  | ||||||||||||||| |||||
Db        411 TGCCCCATGGAGCTGGTTCGCTGTATCCGGCACATTCTATACAATGAACAGAGGCTGGTC 470

Qy        361 CGAGAAGCCAACAATGGTAGCTCTCCAGCTGGAAGCCTTGCTGATGCCATGTCCCAGAAA 420
              |||||||||||||| || |||||||| |||||||| |||||||| ||||||||||||||||
Db        471 CGAGAAGCCAACAACGGAAGCTCTCCGGCTGGAAGTCTTGCTGACGCCATGTCCCAGAAG 530

Qy        421 CACCTCCAGATCAACCAGACGTTTGAGGAGCTGCGACTGGTCACGCAGGACACAGAGAAT 480
              ||||| |||||||||| |||||||||||||||| ||| |||| |||||||| ||||
Db        531 CACCTTCAGATCAACCAAACGTTTGAGGAGCTGCGTCTGATCACACAGGACACGGAGAGC 590

Qy        481 GAGTTAAAAAAGCTGCAGCAGACTCAGGAGTACTTCATCATCCAGTACCAGGAGAGCCTG 540
              ||| | || |||||||||||||||| || ||||||||||||||||||||||||||||||||
Db        591 GAGCTGAAGAAGCTGCAGCAGACCCAAGAGTACTTCATCATCCAGTACCAGGAGAGCCTA 650

Qy        541 AGGATCCAAGCTCAGTTTGGCCCGCTGGCCCAGCTGAGCCCCCAGGAGCGTCTGAGCCGG 600
              |||||||||||||||||| || ||||||||| |||| ||||||||||||  ||||| ||
Db        651 CGGATCCAAGCTCAGTTTGCCCAGCTGGCCCAACTGAACCCCCAGGAGCGCATGAGCAGG 710

Qy        601 GAGACGGCCCTCCAGCAGAAGCAGGTGTCTCTGGAGGCCTGGTTGCAGCGTGAGGCACAG 660
              ||||| |||||||||||||||| ||||| |||||| ||||| ||||||| |||||||||
Db        711 GAGACAGCCCTCCAGCAGAAGCAAGTGTCCCTGGAGACCTGGCTGCAGCGAGAGGCACAG 770

Qy        661 ACACTGCAGCAGTACCGCGTGGAGCTGCCCGAGAAGCACCAGAAGACCCTGCAGCTGCTG 720
              |||||||||||||||||||||||||||||| | || ||||||||||||||||||||||||
Db        771 ACACTGCAGCAGTACCGCGTGGAGCTGGCTGAAAAGCACCAGAAGACCCTGCAGCTGCTG 830

Qy        721 CGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCAGCAG 780
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||| |||||
Db        831 AGGAAGCAGCAGACCATCATCCTGGACGACGAGCTGATCCAGTGGAAGCGGCGACAGCAA 890

Qy        781 CTGGCCGGGAACGGCGGGCCCCCCGAGGGCAGCCTGGACGTGCTACAGTCCTGGTGTGAG 840
              |||||||||||| || || || || |||||||||||||| ||||| ||||||||||||||||
```



Search completed: May 17, 2005, 01:42:31
Job time : 1838 secs